

DYNAMIC DEPENDENCY PAIRS FOR ALGEBRAIC FUNCTIONAL SYSTEMS *

CYNTHIA KOP AND FEMKE VAN RAAMSDONK

Faculty of Sciences, VU University, De Boelelaan 1081a, 1081 HV Amsterdam, The Netherlands
e-mail address: kop@few.vu.nl, femke@few.vu.nl

ABSTRACT. We extend the higher-order termination method of dynamic dependency pairs to Algebraic Functional Systems (AFSs). In this setting, simply typed lambda-terms with algebraic reduction and separate β -steps are considered. For left-linear AFSs, the method is shown to be complete. For so-called local AFSs we define a variation of usable rules and an extension of argument filterings. All these techniques have been implemented in the higher-order termination tool WANDA.

1. INTRODUCTION

An important method to (automatically) prove termination of first-order term rewrite systems is the dependency pair approach by Arts and Giesl [2]. This approach transforms a term rewrite system into groups of ordering constraints, such that rewriting is terminating if and only if the groups of constraints are (separately) solvable. These constraints can be simplified using for instance argument filterings and usable rules [2, 11, 15]. Various optimisations of the method have been studied, see for example [14, 10].

This paper contributes to the study of dependency pairs for higher-order rewriting. It is not easy to adapt the approach to a higher-order setting, primarily due to the presence of β -reduction. A first, very natural extension to Nipkow's HRSs (higher-order rewrite systems) is given in [35], but it relies on the *subterm property*. Due to this property it is impossible to define optimisations like argument filterings. Moreover, unlike the first-order case, the method is not complete: a terminating system may well have an infinite dependency chain.

Since then, the focus of higher-order dependency pairs has been on the so-called *static* style. This style imposes limitations on the rewrite rules which allow the subterm property to be dropped. For static dependency pairs, too, there are no completeness results available.

Here we return to the original, *dynamic* style of dependency pairs, and show how the subterm property can be weakened. We introduce variations of usable rules and argument

1998 ACM Subject Classification: F4.1, F4.2.

Key words and phrases: higher-order rewriting, termination, dynamic dependency pairs.

* Extended version of [24].

This research is supported by the Netherlands Organisation for Scientific Research (NWO-EW) under grant 612.000.629 (Higher-Order Termination).

filterings, and use weakly monotonic algebras and recursive path orderings to orient the resulting constraints. While special interest is reserved for the class of so-called *local* systems, the core technique is defined without restrictions, and is complete for left-linear systems.

Unlike previous approaches, we do not consider rewriting modulo β/η (Nipkow’s HRSs [30]), but with β -reduction as a separate step (Jouannaud and Okada’s AFSs [18]). Although higher-order path orderings are commonly studied in the setting of AFSs [19, 4], and it is the only style of higher-order rewriting which currently appears in the annual termination competition [38], there is so far little work on dependency pairs for this formalism.

This paper is an extended version of [24], with complete proofs and some new features.

Paper Setup We briefly discuss the ideas from studies of dependency pairs for HRSs and also for applicative systems in Section 2. In Section 3 we recapitulate the AFS formalism, and in Section 4 we give a brief overview of the dependency pair framework for first-order term rewriting. Basic (unrestricted) definitions of higher-order dependency pairs, dependency chains, the dependency graph and reduction orderings are discussed in Section 5.

In order to obtain stronger results, we then introduce a restriction “local”, which many common AFSs satisfy, and define *formative rules* (a variation of usable rules) for local AFSs. We show how the results from Section 5 can be strengthened for local systems.

To find reduction orderings for dependency pair constraints, Section 7 discusses two approaches: first, we will see how dependency pairs interact with weakly monotonic algebras, and then we define *argument functions*, a generalisation of argument filterings.

In Section 8 we discuss some improvements when dealing with non-collapsing dependency pairs; for example, in this setting we can use the subterm criterion and usable rules. Section 9 summarises all results, both for local and non-local systems, in a ready-to-use algorithm.

In Section 10 we discuss how the theory in this paper can be used for polymorphic and otherwise infinite systems, and how the static and dynamic dependency pair approaches can be combined. Experimental results with our tool WANDA are presented in Section 11.

2. BACKGROUND AND RELATED WORK

In this section, we discuss the existing work on higher-order dependency pairs.

The existing work on higher-order dependency pairs can roughly be split along two axes. On the one axis, the higher-order formalism; we distinguish between applicative rewriting, rewriting modulo β (HRSs), and with β as a separate step (AFSs). On the other the style of dependency pairs, with common styles being *dynamic* and *static*. Figure 1 gives an overview.

| | Applicative | HRS | AFS |
|----------------|-------------|--------------------|------------------------|
| Dynamic | [25] [1] | [35] [23] | [24] <i>this paper</i> |
| Static | [27] [28] | [3] [34] [26] [36] | [3] |
| Other | [16] | – | – |

Figure 1: *Papers on Higher Order Dependency Pairs*

The dynamic and static approach differ in the treatment of leading variables in the right-hand sides of rules (subterms $x \cdot s_1 \cdots s_n$ with $n > 0$ and x a free variable). In the dynamic approach, such subterms lead to a dependency pair; in the static approach they do not. First-order techniques like argument filterings, the subterm criterion and usable rules are easier to extend to a static approach, while equivalence results tend to be limited to the dynamic style. Static dependency pairs rely on certain restrictions on the rules.

Dependency pairs for applicative term rewriting In applicative systems, terms are built from variables, constants and a binary application operator. Functional variables may be present, as in $x \cdot a$, but there is no abstraction, as in $\lambda x. x$. There are various styles of applicative rewriting, both untyped, simply typed, and with alternative forms of typing.

A dynamic approach was defined both for untyped and simply-typed applicative systems in [25], along with a definition of argument filterings. A first static approach appears in [27] and is improved in [28]; the method is restricted to ‘plain function passing’ systems where, intuitively, leading variables are harmless. Due to the lack of binders, it is also possible to eliminate leading variables by instantiating them, as is done for simply-typed systems in [1]; in [16], an uncurrying transformation from untyped applicative systems to normal first-order systems is used. These techniques have no parallel in rewriting with binders.

Unfortunately, strong though the results for applicative systems may be, they are not directly useful in the setting of AFSs, since termination may be lost by adding λ -abstraction and β -reduction. For example, the simply-typed applicative system $\mathbf{app} \cdot (\mathbf{abs} \cdot F) \cdot x \rightarrow F \cdot x$, with $F : \iota \Rightarrow \iota$ a functional variable, $x : \iota$ a variable, and $\mathbf{app} : \iota \Rightarrow \iota \Rightarrow \iota$, $\mathbf{abs} : (\iota \Rightarrow \iota) \Rightarrow \iota$ constants, is terminating because in every step the size of a term decreases. However, adding λ -abstraction and β -reduction destroys this property: with $\omega = \mathbf{abs} \cdot (\lambda x. \mathbf{app} \cdot x \cdot x)$ we have $\mathbf{app} \cdot \omega \cdot \omega = \mathbf{app} \cdot (\mathbf{abs} \cdot (\lambda x. \mathbf{app} \cdot x \cdot x)) \cdot \omega \rightarrow (\lambda x. \mathbf{app} \cdot x \cdot x) \cdot \omega \rightarrow \mathbf{app} \cdot \omega \cdot \omega$.

Let us move on to rewriting with binders; most results here are on Nipkow’s HRSs.

Dynamic Dependency Pairs for HRSs A first definition of dependency pairs for HRSs is given in [35]. Here termination is not equivalent to the absence of infinite dependency chains, and a term is required to be greater than its subterms (the *subterm property*), which makes many optimisations impossible. In [23] (extended abstract) we have discussed how the subterm property may be weakened by posing restrictions on the rules, and in [24], the short version of this paper, we have explored an extension of the dynamic approach to AFSs.

Static Dependency Pairs for HRSs The static approach in [27] is moved to the setting of HRSs in [26], and extended with argument filterings and usable rules in [36]. The static approach omits dependency pairs $f^\#(\vec{l}) \rightsquigarrow x \cdot \vec{r}$ with x a variable, which avoids the need for a subterm property, but it allows bound variables to become free in the right-hand side of a dependency pair. The technique is restricted to *plain function passing* HRSs. A system with for instance the (terminating) rule $\mathbf{h}(\mathbf{g}(\lambda x. F(x))) \rightarrow F(\mathbf{a})$ cannot be handled. Moreover, the approach is not complete: a terminating AFS may have a static dependency chain.

The definitions for HRSs [35, 26] do not immediately carry over to AFSs, since AFSs may have rules of functional type, and β -reduction is a separate rewrite step. A short paper by Blanqui [3] introduces static dependency pairs on a form of rewriting which includes AFSs, but it poses some restrictions, such as base-type rules. The present work considers dynamic dependency pairs for AFSs and is most related to [35], but is adapted for the different formalism. Our method conservatively extends the one for first-order rewriting and provides a characterisation of termination for left-linear AFSs. We have chosen for a dynamic rather than a static approach because, although the static approach is stronger when applicable, the dynamic definitions can be given without restrictions. The restrictions we do provide, to weaken the subterm property and enable for instance argument filterings, are optional. We will say some words about integrating the static and dynamic approaches in Section 10.

3. PRELIMINARIES

In this section, we present the formalism of Algebraic Functional Systems (AFSs).

We consider higher-order rewriting as defined originally by Jouannaud and Okada [18], also called Algebraic Functional Systems (AFSs). Terms are built from simply-typed variables, abstraction and application (as in the simply-typed λ -calculus), and also from function symbols which take a fixed number of arguments. Terms and matching are modulo α , and every AFS contains the β -reduction rule. Several variations of the definition of AFSs exist; here we roughly follow [37, Chapter 11.2.3], which coincides with the format currently used in the higher-order category of the annual termination competition [38].

Types and Terms Assuming a set \mathcal{B} of *base types*, the set of *simple types* (or just *types*) is generated using the binary type constructor \Rightarrow , according to the following grammar:

$$\mathbb{T} ::= \mathcal{B} \mid \mathbb{T} \Rightarrow \mathbb{T}$$

The arrow operator is right-associative. Types are denoted by σ, τ, \dots and base types by ι, κ, \dots . A type with at least one occurrence of \Rightarrow is called a *functional type*. A *type declaration* is an expression of the form $[\sigma_1 \times \dots \times \sigma_n] \Rightarrow \tau$ with τ and all σ_i types; we write just τ if $n = 0$. Type declarations are not types, but are used for typing purposes.

We assume a set \mathcal{V} , consisting of for each type infinitely many typed *variables*, written as x, y, z, \dots . We further assume a set \mathcal{F} , disjoint from \mathcal{V} , consisting of *function symbols*, equipped with a type declaration, and written as f, g, \dots or using more suggestive notation. To stress the type (declaration) of a symbol a we may write $a : \sigma$. The set of *terms* over \mathcal{F} consists of expressions s for which we can infer $s : \sigma$ for some type σ using the clauses:

$$\begin{array}{lll} (\text{var}) & x : \sigma & \text{if } x : \sigma \in \mathcal{V} \\ (\text{app}) & s \cdot t : \tau & \text{if } s : \sigma \Rightarrow \tau \text{ and } t : \sigma \\ (\text{abs}) & \lambda x. s : \sigma \Rightarrow \tau & \text{if } x : \sigma \in \mathcal{V} \text{ and } s : \tau \\ (\text{fun}) & f(s_1, \dots, s_n) : \tau & \text{if } f : [\sigma_1 \times \dots \times \sigma_n] \Rightarrow \tau \in \mathcal{F} \text{ and } s_1 : \sigma_1, \dots, s_n : \sigma_n \end{array}$$

Terms built using these clauses are called respectively a variable, an application, an abstraction, and a functional term. Note that a function symbol $f : [\sigma_1 \times \dots \times \sigma_n] \Rightarrow \tau$ takes exactly n arguments, and τ may be a functional type. The λ binds occurrences of variables as in the λ -calculus, and term equality is modulo α -conversion (bound variables may be renamed). A variable in s which is not bound by some λ is *free*, and the set of free variables of s is denoted by $FV(s)$. Application is left-associative. Let $\text{head}(\cdot)$ denote the head of an application, so $\text{head}(s \cdot t) = \text{head}(s)$ and $\text{head}(s) = s$ for non-applications.

A *substitution* $[\vec{x} := \vec{s}]$, with \vec{x} and \vec{s} non-empty finite vectors of equal length, is the homomorphic extension of the type-preserving mapping $\vec{x} \mapsto \vec{s}$ from variables to terms. Substitutions are denoted γ, δ , and the result of applying γ to a term s is denoted $s\gamma$. The *domain* $\text{dom}(\gamma)$ of $\gamma = [\vec{x} := \vec{s}]$ is $\{\vec{x}\}$. Substituting does not capture free variables.

Let $\square_\sigma : \sigma$ be a fresh symbol for every type σ . A *context* $C[\cdot]$ is a term with a single occurrence of some \square_σ . The result of the replacement of \square_σ in $C[\cdot]$ by a term $s : \sigma$ is denoted by $C[s]$. Such replacements may capture free variables. For example, $(\lambda x. y)[y := x] = \lambda z. x$, but for $C[\cdot] = \lambda x. \square_\sigma$ we have $C[x] = \lambda x. x$.

We say t is a *subterm* of s , notation $s \triangleright t$, if $s = C[t]$ for some context C . If in addition $C \neq \square_\sigma$, then t is a *strict subterm* of s , notation $s \triangleright t$.

Rules and Rewriting A *rewrite rule* over a set of function symbols \mathcal{F} is a pair of terms $l \rightarrow r$ over \mathcal{F} such that l and r have the same type, and all free variables of r also occur in

l . In [21] some termination-preserving transformations on the general format of AFS-rules are presented. Using these results, we can additionally assume that a left-hand side l is of the form $f(l_1, \dots, l_n) \cdot l_{n+1} \cdots l_m$ (with $m \geq n \geq 0$), and does not contain subterms of the form $(\lambda x. s) \cdot t$. (Many AFSs are defined in this way already.) Note that we do not assume η -normal or η -exanded forms, and that we may have rules of functional type.

Given a set of rewrite rules \mathcal{R} , the *rewrite relation* $\rightarrow_{\mathcal{R}}$ on terms is given by:

$$\begin{aligned} \text{(rule)} \quad & C[l\gamma] \rightarrow_{\mathcal{R}} C[r\gamma] \quad \text{with } l \rightarrow r \in \mathcal{R}, C \text{ a context, } \gamma \text{ a substitution} \\ \text{(beta)} \quad & C[(\lambda x. s) \cdot t] \rightarrow_{\mathcal{R}} C[s[x := t]] \end{aligned}$$

We sometimes use the notation $s \rightarrow_{\beta} t$ for a rewrite step using **(beta)**. A *headmost step* is a reduction $s \rightarrow_{\mathcal{R}} t$ using either clause, where C has the form $\square_{\sigma} \cdot s_1 \cdots s_n$ with $n \geq 0$.

An *algebraic functional system* (AFS) is a pair $(\mathcal{F}, \mathcal{R})$ consisting of a set of function symbols \mathcal{F} and a set \mathcal{R} of rewrite rules over \mathcal{F} ; it is often specified by giving only \mathcal{R} .

A function symbol f is a *defined symbol* of an AFS if there is a rule with left-hand side $f(l_1, \dots, l_n) \cdot l_{n+1} \cdots l_m$, and a *constructor symbol* if not. The sets of defined and constructor symbols are denoted by \mathcal{D} and \mathcal{C} respectively. A rewrite rule $l \rightarrow r$ is *left-linear* if every variable occurs at most once free in l ; an AFS is left-linear if all its rewrite rules are. A rule $l \rightarrow r$ is *collapsing* if $\text{head}(r)$ is a variable. A term s is called *terminating* if every reduction sequence starting in s is finite; an AFS is terminating if all its terms are.

We assume that an AFS has only finitely many rules. In Section 10.2 we shortly discuss how to use dependency pairs to prove termination of AFSs with infinitely many rules.

Example 3.1. The following AFS *twice* is the running example of this paper. It has the following four function symbols: $\circ : \text{nat}$, $\mathsf{s} : [\text{nat}] \Rightarrow \text{nat}$, $\mathsf{l} : [\text{nat}] \Rightarrow \text{nat}$, and $\text{twice} : [\text{nat} \Rightarrow \text{nat}] \Rightarrow \text{nat} \Rightarrow \text{nat}$. There are three rewrite rules:

$$\begin{aligned} \mathsf{l}(\circ) &\rightarrow \circ \\ \mathsf{l}(\mathsf{s}(n)) &\rightarrow \mathsf{s}(\text{twice}(\lambda x. \mathsf{l}(x)) \cdot n) \\ \text{twice}(F) &\rightarrow \lambda y. F \cdot (F \cdot y) \end{aligned}$$

Recall that we also have β -reduction steps. The symbol l represents the identity function on natural numbers. This system is terminating, but this is not trivial to prove; neither recursive path orderings like HORPO [19] and CPO [4], nor a static dependency pair approach, can handle the second l -rule, due to the subterm $\mathsf{l}(x)$. The static approach gives a constraint $\mathsf{l}^\sharp(\mathsf{s}(n)) \succ \mathsf{l}^\sharp(x)$, which is impossible to satisfy because \succ must be closed under substitution, and $\mathsf{s}(n)$ might be substituted for the free variable x . CPO gives a similar problem.

4. THE FIRST-ORDER DEPENDENCY PAIR APPROACH

In this section, we recall the dependency pair approach for first-order rewriting. We emphasise those parts which are relevant for our higher-order approach.

We assume that first-order term rewrite systems (TRSs) are already known; they can also be thought of as AFSs where all function symbols have a type declaration $[\circ \times \dots \times \circ] \Rightarrow \circ$ and where terms are formed without clauses **(abs)** and **(app)**. In this section we recall those definitions and results from the theory of dependency pairs for TRSs that we will generalize or adapt to the higher-order setting in this paper. The definitions here are close to those in [13]; our set-up is in between the one for the dependency pairs *approach* [2] and the dependency pairs *framework* [10]. This section is also meant to give additional background for those not familiar with dependency pairs.

4.1. Motivation. Two important properties of the termination method using dependency pairs are that it is suitable for automation, and that it can be used to prove termination of TRSs which are not simply terminating. A (well-known) example of such a TRS is:

$$\begin{array}{llll} \text{minus}(x, \mathbf{o}) & \rightarrow & x & \text{quot}(\mathbf{o}, \mathbf{s}(y)) \rightarrow \mathbf{o} \\ \text{minus}(\mathbf{s}(x), \mathbf{s}(y)) & \rightarrow & \text{minus}(x, y) & \text{quot}(\mathbf{s}(x), \mathbf{s}(y)) \rightarrow \mathbf{s}(\text{quot}(\text{minus}(x, y), \mathbf{s}(y))) \end{array}$$

A TRS is simply terminating if it can be proved terminating using a *reduction ordering* \succ (a well-founded ordering on terms which is both monotonic and closed under substitution) that satisfies the *subterm property*, which means that $f(s_1, \dots, s_n) \succeq s_i$ for every $i \in \{1, \dots, n\}$.

4.2. Dependency Pairs. An intuition behind the dependency pair approach is to identify those parts of the right-hand sides of rewrite rules which may give rise to an infinite reduction. Suppose we have a minimal non-terminating term t , so a non-terminating term where all proper subterms are terminating. An infinite reduction from t has the form $t \rightarrow^* l\gamma \rightarrow r\gamma \rightarrow \dots$ with $l \rightarrow r$ a rewrite rule. Then, a minimal non-terminating subterm of $r\gamma$ has as root-symbol a defined symbol from the pattern of r . Thus, we are interested in subterms of right-hand sides of rewrite rules with a defined symbol at the root. Such subterms are called *candidate terms* of r .

We obtain $\mathcal{F}^\#$ by adding to each symbol f in the signature \mathcal{F} of a first-order TRS a symbol $f^\#$ with the same arity. The *dependency pairs* of a rewrite rule $f(l_1, \dots, l_n) \rightarrow r$ are all pairs $f^\#(l_1, \dots, l_n) \rightsquigarrow g^\#(p_1, \dots, p_m)$ with $r \geq g(p_1, \dots, p_m)$, and g a defined symbol, and $g(p_1, \dots, p_m)$ not a subterm of some l_i . The set of all dependency pairs of a TRS $(\mathcal{F}, \mathcal{R})$ is denoted by $\text{DP}(\mathcal{R})$.

The quot -example has the following dependency pairs:

$$\begin{array}{ll} \text{minus}^\#(\mathbf{s}(x), \mathbf{s}(y)) & \rightsquigarrow \text{minus}^\#(x, y) \\ \text{quot}^\#(\mathbf{s}(x), \mathbf{s}(y)) & \rightsquigarrow \text{quot}^\#(\text{minus}(x, y), \mathbf{s}(y)) \\ \text{quot}^\#(\mathbf{s}(x), \mathbf{s}(y)) & \rightsquigarrow \text{minus}^\#(x, y) \end{array}$$

The first and third rewrite rule do not give dependency pairs, because their right-hand sides do not contain defined symbols. The fourth rule gives two different dependency pairs.

A *dependency chain* is a sequence $[(l_i \rightsquigarrow p_i, s_i, t_i) \mid i \in \mathbb{N}]$, such that for all i :

- (1) $l_i \rightsquigarrow p_i \in \text{DP}(\mathcal{R})$;
- (2) $s_i = l_i\gamma_i$ and $t_i = p_i\gamma_i$ for some substitution γ_i ;
- (3) $t_i \rightarrow_{\mathcal{R}}^* s_{i+1}$.

Since t_i has the form $f^\#(\vec{u})$ and the marked symbol $f^\#$ is not used in any rule, the reduction $t_i \rightarrow_{\mathcal{R}}^* s_{i+1}$ does not use headmost steps. The chain is *minimal* if all t_i are terminating under $\rightarrow_{\mathcal{R}}$. Termination of a TRS can be characterized using dependency chains:

Theorem 4.1 ([2]). *A TRS is terminating if and only if it does not admit a minimal dependency chain.*

A higher-order generalization of this result is provided in Theorems 5.7 and 5.9.

Example 4.2. The TRS $\text{nats}(n) \rightarrow \text{cons}(n, \text{nats}(\mathbf{s}(n)))$ has an infinite dependency chain

$$\begin{array}{l} (\text{nats}^\#(n) \rightsquigarrow \text{nats}^\#(\mathbf{s}(n)) \quad , \quad \text{nats}^\#(\mathbf{o}) \quad , \quad \text{nats}^\#(\mathbf{s}(\mathbf{o})) \quad), \\ (\text{nats}^\#(n) \rightsquigarrow \text{nats}^\#(\mathbf{s}(n)) \quad , \quad \text{nats}^\#(\mathbf{s}(\mathbf{o})) \quad , \quad \text{nats}^\#(\mathbf{s}(\mathbf{s}(\mathbf{o}))) \quad), \\ \dots \end{array}$$

corresponding to the infinite reduction $\text{nats}(\mathbf{o}) \rightarrow \text{nats}(\mathbf{s}(\mathbf{o})) \rightarrow \text{nats}(\mathbf{s}(\mathbf{s}(\mathbf{o}))) \rightarrow \dots$

4.3. Using a Reduction Pair. A TRS without infinite dependency chain is terminating. Absence of infinite dependency chains can be demonstrated with a *reduction pair*. This is a pair (\succeq, \succ) of a *quasi-order* and a *well-founded order*, such that:

- \succeq and \succ are *compatible*: either $\succ \cdot \succeq$ is included in \succ , or $\succeq \cdot \succ$ is;
- \succeq and \succ are both *stable* (preserved under substitution);
- \succeq is monotonic (if $s \succeq t$, then $C[s] \succeq C[t]$).

Theorem 4.3 ([2]). *A TRS is terminating, if there exists a reduction pair (\succeq, \succ) such that $l \succ p$ for all dependency pairs $l \rightsquigarrow p$, and $l \succeq r$ for all rules $l \rightarrow r$.*

Example 4.4. The `quot` example is terminating if there is a reduction pair satisfying:

$$\begin{array}{llll} \text{minus}(x, o) & \succeq & x & \text{minus}^\#(s(x), s(y)) & \succ & \text{minus}^\#(x, y) \\ \text{minus}(s(x), s(y)) & \succeq & \text{minus}(x, y) & \text{quot}^\#(s(x), s(y)) & \succ & \text{quot}^\#(\text{minus}(x, y), s(y)) \\ \text{quot}(o, s(y)) & \succeq & o & \text{quot}^\#(s(x), s(y)) & \succ & \text{minus}^\#(x, y) \\ \text{quot}(s(x), s(y)) & \succeq & s(\text{quot}(\text{minus}(x, y), s(y))) \end{array}$$

These constraints are oriented with a polynomial interpretation with $\llbracket \text{minus}(x, y) \rrbracket = \llbracket \text{minus}^\#(x, y) \rrbracket = \llbracket x \rrbracket$, $\llbracket \text{quot}(x, y) \rrbracket = \llbracket \text{quot}^\#(x, y) \rrbracket = \llbracket x \rrbracket + \llbracket y \rrbracket$, $\llbracket s(x) \rrbracket = \llbracket x \rrbracket + 1$, $\llbracket o \rrbracket = 0$:

$$\begin{array}{ll} x & \geq x \\ x + 1 & \geq x \\ y + 1 & \geq 0 \\ x + y + 2 & \geq x + y + 2 \end{array} \quad \begin{array}{ll} x + 1 & > x \\ x + y + 2 & > x + y + 1 \\ x + y + 2 & > x \end{array}$$

Note that we needed dependency pairs to use an interpretation like this: the resulting ordering \succ is not monotonic, since $\llbracket \text{minus}(s, t) \rrbracket = \llbracket \text{minus}(s, u) \rrbracket$ even if $t \succ u$.

Rather than orienting all dependency pairs at once, we can use a step by step approach. Let a set of dependency pairs \mathcal{P} be called *chain-free* if there is no minimal dependency chain using only pairs in \mathcal{P}^1 . Note that \emptyset is chain-free. The following result has Theorem 5.16 as a higher-order counterpart.

Theorem 4.5. *A set $\mathcal{P} = \mathcal{P}_1 \uplus \mathcal{P}_2$ is chain-free if \mathcal{P}_2 is chain-free, and there is a reduction pair (\succeq, \succ) such that $l \succ p$ for $l \rightsquigarrow p \in \mathcal{P}_1$, $l \succeq p$ for $l \rightsquigarrow p \in \mathcal{P}_2$ and $l \succeq r$ for $l \rightarrow r \in \mathcal{R}$.*

4.4. Argument Filterings. In order to obtain a reduction pair, there are two ways we could go: either we use approaches like the polynomial interpretations given in Example 4.4, which directly give us a pair (\succeq, \succ) (where \succ may be non-monotonic), or we use an existing reduction ordering or pair and adapt it with *argument filterings*. An argument filtering is a function π which maps terms of the form $f(x_1, \dots, x_n)$ with $f \in \mathcal{F}^\#$ either to a term $f_\pi(x_{i_1}, \dots, x_{i_m})$ or to one of the x_i . An argument filtering is applied to a term as follows:

$$\begin{array}{ll} \bar{\pi}(f(s_1, \dots, s_n)) & = f_\pi(\bar{\pi}(s_{i_1}), \dots, \bar{\pi}(s_{i_m})) \quad \text{if } \pi(f(\vec{x})) = f_\pi(x_{i_1}, \dots, x_{i_m}) \\ \bar{\pi}(f(s_1, \dots, s_n)) & = \bar{\pi}(s_i) \quad \text{if } \pi(f(\vec{x})) = x_i \\ \bar{\pi}(x) & = x \quad \text{if } x \text{ a variable} \end{array}$$

Phrased differently, $\bar{\pi}(f(s_1, \dots, s_n)) = \pi(f(x_1, \dots, x_n))[x_1 := \bar{\pi}(s_1), \dots, x_n := \bar{\pi}(s_n)]$. Note that an argument filtering works both on unmarked and on marked symbols.

Using argument filterings, we can eliminate troublesome subterms of the dependency pair constraints. To this end, we use the following result, which corresponds to Theorem 7.8:

¹In the language of [10], this corresponds to finiteness of the DP problem $(\mathcal{P}, \emptyset, \mathcal{R}, \text{minimal})$.

Theorem 4.6. *Given a reduction pair $(\geq, >)$ and an argument filtering π , define: $s \succeq t$ iff $\pi(s) \geq \pi(t)$ and $s \succ t$ iff $\pi(s) > \pi(t)$. Then (\succeq, \succ) is a reduction pair.*

Example 4.7. Recall the constraints given in Example 4.4. We use the argument filtering with $\pi(\text{minus}(x, y)) = x$, $\pi(\text{quot}(x, y)) = \text{quot}_\pi(x)$ and $\pi(f(\vec{x})) = f_\pi(\vec{x})$ for all other symbols. By Theorem 4.6 it suffices to find a reduction ordering satisfying:

$$\begin{array}{ll} x & \geq x & \text{minus}_\pi^\#(s_\pi(x), s_\pi(y)) & > \text{minus}_\pi^\#(x, y) \\ s_\pi(x) & \geq x & \text{quot}_\pi^\#(s_\pi(x), s_\pi(y)) & > \text{quot}_\pi^\#(x, s_\pi(y)) \\ \text{quot}_\pi(o_\pi) & \geq o_\pi & \text{quot}_\pi^\#(s_\pi(x), s_\pi(y)) & > \text{minus}_\pi^\#(x, y) \\ \text{quot}_\pi(s_\pi(x)) & \geq s_\pi(\text{quot}_\pi(x)) \end{array}$$

These altered constraints can easily be satisfied with a lexicographic path ordering.

4.5. The Subterm Criterion. An alternative to reduction pairs, which often suffices to eliminate some dependency pairs and is typically easy to check, is the *subterm criterion*.

A function ν that assigns to every n -ary dependency pair symbol $f^\#$ one of its argument positions is said to be a *projection function*. We extend ν to a function on terms by defining:

$$\bar{\nu}(f^\#(s_1, \dots, s_n)) = s_i \text{ if } \nu(f^\#) = i$$

Theorem 4.8. *A set of dependency pairs $\mathcal{P} = \mathcal{P}_1 \uplus \mathcal{P}_2$ is chain-free if \mathcal{P}_2 is chain-free, and moreover there is a projection function ν such that $\bar{\nu}(l) \triangleright \bar{\nu}(p)$ for all dependency pairs $l \rightsquigarrow p \in \mathcal{P}_1$ and $\bar{\nu}(l) = \bar{\nu}(p)$ for all $l \rightsquigarrow p \in \mathcal{P}_2$.*

This holds because for all s_i and t_i in a dependency chain, $\bar{\nu}(s_i)$ and $\bar{\nu}(t_i)$ are strict subterms, and therefore terminating under $\rightarrow_{\mathcal{R}} \cup \triangleright$; see also Theorem 8.2.

The subterm criterion is not sufficient to show termination of the `quot` example, but at least we can use it to eliminate some dependency pairs; choosing $\nu(\text{minus}^\#) = \nu(\text{quot}^\#) = 2$:

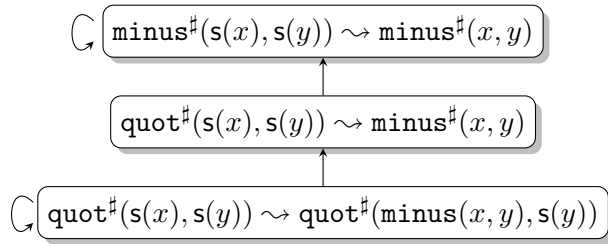
$$\begin{array}{llll} \bar{\nu}(\text{minus}^\#(s(x), s(y))) & = s(y) & \triangleright y & = \bar{\nu}(\text{minus}^\#(x, y)) \\ \bar{\nu}(\text{quot}^\#(s(x), s(y))) & = s(y) & = s(y) & = \bar{\nu}(\text{quot}^\#(\text{minus}(x, y), s(y))) \\ \bar{\nu}(\text{quot}^\#(s(x), s(y))) & = s(y) & \triangleright y & = \bar{\nu}(\text{minus}^\#(x, y)) \end{array}$$

This shows that the TRS `quot` is non-terminating if and only if there is no dependency chain where every step uses the dependency pair $\text{quot}^\#(s(x), s(y)) \rightsquigarrow \text{quot}^\#(\text{minus}(x, y), s(y))$.

4.6. The Dependency Graph. To determine whether a system has a dependency chain, it makes sense to ask what form such a chain would have. This question is studied with a *dependency graph*, a graph with as nodes the dependency pairs of \mathcal{R} and an edge from $l \rightsquigarrow p$ to $u \rightsquigarrow v$ if $p\gamma \rightarrow_{\mathcal{R}}^* u\delta$ for some substitutions γ, δ . See for example the dependency graph of the `quot`-TRS.

If there is a dependency chain $[(\rho_i, s_i, t_i) \mid i \in \mathbb{N}]$, then there is an edge in the graph from each ρ_i to ρ_{i+1} . Since the graph is finite, a dependency chain corresponds to a cycle in the graph.

By definition, if a set of dependency pairs is chain-free, then the same holds for any subset. Since a dependency graph might have exponentially many cycles, modern approaches typically consider only *maximal cycles*, also called *strongly connected components* (SCCs).



Theorem 4.9. *\mathcal{R} is terminating iff every SCC of its dependency graph is chain-free.*

This result is extended to AFSs in Lemma 5.12. The dependency graph is not in general computable, which is why *approximations* are often used. An approximation of a dependency graph G is a graph with the same nodes as G , but which may have additional edges.

The dependency graph of our running example `quot` has two cycles. In order to prove termination, it is sufficient to find a reduction pair such that $\text{minus}^\#(s(x), s(y)) \succ \text{minus}^\#(x, y)$ and $l \succeq r$ for all rules, and a (nother) reduction pair with $\text{quot}^\#(s(x), s(y)) \succ \text{quot}^\#(\text{minus}(x, y), s(y))$ and $l \succeq r$ for all rules. The fact that we can deal with groups of dependency pairs separately can make it significantly simpler to find reduction pairs.

Extending Theorem 4.9, we can iterate over a dependency graph approximation, and obtain the following algorithm (whose higher-order counterpart is presented in Section 9):

Theorem 4.10 ([14]). *A TRS is terminating if and only if this can be demonstrated with the following algorithm:*

- (1) *calculate dependency pairs, find an approximation G for the dependency graph;*
- (2) *if G has no cycles, the TRS is terminating; otherwise, choose an SCC \mathcal{P}*
- (3) *try finding a projection function $\bar{\nu}$ such that $\bar{\nu}(l) \triangleright \bar{\nu}(p)$ for at least some $l \rightsquigarrow p \in \mathcal{P}$ and $\bar{\nu}(l) \sqsupseteq \bar{\nu}(p)$ for the rest; if this succeeds, remove the strictly oriented pairs from G and continue with 2.*
- (4) *find a reduction pair (\succeq, \succ) such that $l \succeq r$ for all rules $l \rightarrow r$, and for all dependency pairs $l \rightsquigarrow p \in \mathcal{P}$ either $l \succ p$ or $l \succeq p$; at least one pair must be oriented with \succ (**);*
- (5) *remove the dependency pairs which were oriented with \succ from G ; continue with 2.*
- (**) *To find (\succeq, \succ) we may for instance use argument filterings.*

4.7. Usable Rules. We discuss one more optimisation. In the algorithm of Theorem 4.10, we consider in every iteration for a strongly connected component \mathcal{P} all rewrite rules. Instead, we can restrict attention to the rules that may be relevant for constructing a dependency chain using dependency pairs from \mathcal{P} . To this end the concept of *usable rules* is defined.

First we need some definitions. We denote by $f \sqsubseteq_{us} g$ that there is a rewrite rule $f(l_1, \dots, l_n) \rightarrow C[g(r_1, \dots, r_m)]$. The reflexive-transitive closure of \sqsubseteq_{us} is denoted by \sqsubseteq_{us}^* . Overloading notation, we write $s \sqsubseteq_{us}^* g$ if there is a symbol f in the term s such that $f \sqsubseteq_{us}^* g$. So if not $s \sqsubseteq_{us}^* g$, then s cannot reduce to a term containing the symbol g .

Definition 4.11. The set of *usable rules* of a term s , notation $UR(s)$, consists of rules $g(\vec{l}) \rightarrow r \in \mathcal{R}$, where $s \sqsubseteq_{us}^* g$. For a set of dependency pairs \mathcal{P} , let $UR(\mathcal{P}) = \bigcup_{l \rightsquigarrow p \in \mathcal{P}} UR(p)$.

Using a reasoning originally due to Gramlich [12], it is shown in [13] that if \mathcal{P} is not chain-free, then there is a dependency chain over \mathcal{P} where the reduction $t_i \rightarrow^* s_{i+1}$ uses only the rules in $UR(\mathcal{P}) \cup \{\mathbf{p}(x, y) \rightarrow x, \mathbf{p}(x, y) \rightarrow y\}$ for some fresh symbol \mathbf{p} (these two rules are usually considered harmless). Thus, in each iteration of the algorithm of Theorem 4.10 we only have to prove $l \succeq r$ for the usable rules of \mathcal{P} , rather than for all rules.

5. THE BASIC HIGHER-ORDER DEPENDENCY PAIR APPROACH

In this section we define a basic dependency pair approach for AFSs. We show that an AFS is terminating if it does not have a minimal dependency chain, and that for left-linear AFSs, the absence of (minimal) dependency chains characterises termination. As in the first-order case, we organise the dependency pairs in a graph, and explain how to use reduction pairs.

When extending the first-order dependency pair approach to AFSs, new issues arise:

- *collapsing rules*: non-termination might also be caused by a higher-order variable being instantiated. For example, the right-hand side of the non-terminating rule $f(g(F), x) \rightarrow F \cdot x$ doesn't even have defined symbols;
- *dangling variables*: given a rule $f(o) \rightarrow g(\lambda x. f(x))$, the bound variable x should probably not become free in the corresponding dependency pair;
- *rules of functional type* may lead to non-termination only because of their interaction with the (applicative) context they appear in;
- *typing issues*: to be able to use the usual term orderings, both sides of a dependency pair (or the constraints generated from it) should have, usually, the same type modulo renamings of base types.

Typing issues will be addressed in Section 5.5; for the other problems we have to take precautions already in the definition of dependency pairs.

5.1. Dependency Pairs. In order to define dependency pairs, we first pre-process the rewrite rules and define candidate terms. The complete definition of dependency pairs may at first seem somewhat baroque; this is partly because we have to work around the issues of functional rules and dangling variables, and partly because of several optimisations we include to obtain an easier result system.

Pre-processing Pre-processing the rewrite rules is done by completion:

Definition 5.1 (Pre-processing Rules). An AFS is *completed* by adding for each rule of the form $l \rightarrow \lambda x_1 \dots x_n. r$, with r not an abstraction, the following n rewrite rules: $l \cdot x_1 \rightarrow \lambda x_2 \dots x_n. r$, \dots , $l \cdot x_1 \dots x_n \rightarrow r$.

Note that completing an AFS has no effect on termination, since the added rules can be simulated by at most $n + 1$ steps using only the original rules.

Example 5.2. The system twice from Example 3.1 is completed by adding the rewrite rule $\text{twice}(F) \cdot n \rightarrow F \cdot (F \cdot n)$.

In the remainder of the paper, we assume that all AFSs are completed.

To understand why completion is necessary, consider the AFS with a single rule $f(o) \rightarrow \lambda x. f(x) \cdot x$. The term $f(o)$ in this AFS is terminating, but there *is* an infinite reduction $f(o) \cdot o \rightarrow (\lambda x. f(x) \cdot x) \cdot o \rightarrow_\beta f(o) \cdot o \rightarrow \dots$. Rules like this might complicate the analysis of dependency chains, because the important step does not happen at the top. The pre-processing makes sure that it could also be done with a topmost step: $f(o) \cdot o$ self-reduces with a single step using the new rule $f(o) \cdot x \rightarrow f(x) \cdot x$ which was added by completion.

It is worth noting that we did not add new rules for all functional rules, only for those where the right-hand side is an abstraction. A rule $f(o) \rightarrow f(A)$ of functional type is left alone. This is an optimisation: it would be natural to add a rule $f(o) \cdot x \rightarrow f(A) \cdot x$, but this might give a dependency pair $f(o) \cdot x \rightsquigarrow A^\sharp$ which won't be needed. Instead of completing this rule, we will later add a special dependency pair for it.

Candidate terms In the first-order definition of dependency pairs, we identify subterms that may give rise to an infinite reduction. Taking subterms in a system with binders is well-known to be problematic because bound variables may become free. One solution is to

substitute fresh constants in the place of a bound variable which would otherwise become free. In this way, $F \cdot c$ is a “subterm” of $\lambda x. F \cdot (F \cdot x)$. This is the approach we take here.

We assume for every type σ a fresh symbol $c_\sigma : \sigma$. Sometimes the sub-script indicating the type is omitted. The set of all those symbols is denoted by \mathbb{C} . The symbols c_σ are used to replace bound variables which become free by taking a subterm.

Definition 5.3 (Candidate Terms). Let r be a term in an AFS. A subterm t of r is a *candidate term* of r if either $t = f(t_1, \dots, t_m) \cdot t_{m+1} \cdots t_n$ with f a defined symbol and $n \geq m \geq 0$, or $t = x \cdot t_1 \cdots t_n$ with x free in r and $n > 0$.

If t is a candidate term of r , and $\{x_1 : \sigma_1, \dots, x_n : \sigma_n\}$ is the set of variables which occur bound in r but free in t , then $t[x_1 := c_{\sigma_1}, \dots, x_n := c_{\sigma_n}]$ is a *closed candidate term* of r . We denote the set of closed candidate terms of r by $Cand(r)$.

In the AFS *twice* we have $Cand(F \cdot (F \cdot m)) = \{F \cdot (F \cdot m), F \cdot m\}$ and $Cand(s(\text{twice}(\lambda x. l(x)) \cdot n)) = \{\text{twice}(\lambda x. l(x)) \cdot n, \text{twice}(\lambda x. l(x)), l(c_{\text{nat}})\}$. If f is a defined symbol, then the candidate terms of $f(a) \cdot b \cdot c \cdot d$ are $f(a)$, $f(a) \cdot b$, $f(a) \cdot b \cdot c$ and $f(a) \cdot b \cdot c \cdot d$. Note that for example $x \cdot y$ is not a candidate term of $g(\lambda x. x \cdot y)$ because x occurs only bound.

Dependency Pairs As in the first-order case, the definition of dependency pair uses marked function symbols. Let $\mathcal{F}^\# = \mathcal{F} \cup \{f^\# : \sigma \mid f : \sigma \in \mathcal{D}\}$, so \mathcal{F} extended with for every defined symbol f a marked version $f^\#$ with the same type declaration. We denote by $\mathcal{F}_c^\#$ the union of $\mathcal{F}^\#$ and \mathbb{C} . The marked counterpart of a term s , notation $s^\#$, is $f^\#(s_1, \dots, s_n)$ if $s = f(s_1, \dots, s_n)$ with f in \mathcal{D} , and just s otherwise. For example, $(\text{twice}(F))^\# = \text{twice}^\#(F)$ and $(\text{twice}(F) \cdot m)^\# = \text{twice}(F) \cdot m$. Applications are not marked.

Definition 5.4 (Dependency Pair). The set of *dependency pairs* of a rewrite rule $l \rightarrow r$, notation $DP(l \rightarrow r)$, consists of:

- all pairs $l^\# \rightsquigarrow p^\#$ with $p \in Cand(r)$ such that p is no strict subterm of l ;
- if l has a functional type $\sigma_1 \Rightarrow \dots \Rightarrow \sigma_n \Rightarrow \iota$ ($n \geq 1$) and $head(r)$ is either a variable or a term $f(\vec{s})$ with $f \in \mathcal{D}$: all pairs $l \cdot y_1 \cdots y_k \rightsquigarrow r \cdot y_1 \cdots y_k$ with $k \in \{1, \dots, n\}$ and all y_i are fresh variables.

We use $DP(\mathcal{R})$ (or just DP if \mathcal{R} is clear from context) for the set of all dependency pairs of rewrite rules of an AFS \mathcal{R} .

Example 5.5. The set of dependency pairs of the AFS *twice* consists of:

$$\begin{array}{llll} l^\#(s(n)) \rightsquigarrow \text{twice}(\lambda x. l(x)) \cdot n & \text{twice}^\#(F) \rightsquigarrow F \cdot (F \cdot c_{\text{nat}}) \\ l^\#(s(n)) \rightsquigarrow \text{twice}^\#(\lambda x. l(x)) & \text{twice}^\#(F) \rightsquigarrow F \cdot c_{\text{nat}} \\ l^\#(s(n)) \rightsquigarrow l^\#(c_{\text{nat}}) & \text{twice}(F) \cdot m \rightsquigarrow F \cdot (F \cdot m) \\ & \text{twice}(F) \cdot m \rightsquigarrow F \cdot m \end{array}$$

The last two dependency pairs originate from the rule added by completion.

The second form of dependency pair deals with functional rules whose right-hand side is not an abstraction. To illustrate why they are necessary, consider the system with function symbols $A : [\circ] \Rightarrow \circ \Rightarrow \circ$ and $B : [\circ \Rightarrow \circ] \Rightarrow \circ$, and one rewrite rule: $A(B(F)) \rightarrow F$. This system has no dependency pairs of the first kind, but does admit a two-step loop: $s := A(B(\lambda x. A(x) \cdot x)) \cdot B(\lambda x. A(x) \cdot x) \rightarrow (\lambda x. A(x) \cdot x) \cdot B(\lambda x. A(x) \cdot x) \rightarrow_\beta s$. The rule *does* have a dependency pair of the second form, $A(B(F)) \cdot x \rightsquigarrow F \cdot x$.

Comparing our approach to static dependency pairs as defined in [26], the two main differences are that we avoid bound variables becoming free, and that we include *collapsing* dependency pairs, where the right-hand side is headed by a variable.

5.2. Dependency Chains. We can now investigate termination using *dependency chains*:

Definition 5.6. A *dependency chain* is an infinite sequence $[(\rho_i, s_i, t_i) \mid i \in \mathbb{N}]$ such that for all i :

- (1) $\rho_i \in \text{DP} \cup \{\text{beta}\}$;
- (2) if $\rho_i = l_i \rightsquigarrow p_i \in \text{DP}$ then there exists a substitution γ such that $s_i = l_i\gamma$ and $t_i = p_i\gamma$;
- (3) if $\rho_i = \text{beta}$ then $s_i = (\lambda x. u) \cdot v \cdot w_1 \cdots w_k$ and either
 - (a) $k > 0$ and $t_i = u[x := v] \cdot w_1 \cdots w_k$, or
 - (b) $k = 0$ and there exists a term w such that $u \sqsupseteq w$ and $x \in FV(w)$ and $t_i = w^\sharp[x := v]$, but $w \neq x$;
- (4) $t_i \rightarrow_{in}^* s_{i+1}$.

A step \rightarrow_{in} is obtained by rewriting some u_i inside a term of the form $f(u_1, \dots, u_n) \cdot u_{n+1} \cdots u_m$. If $t_i = s_{i+1}$, then also $t_i \rightarrow_{in}^* s_{i+1}$, regardless of whether t_i has this form. A dependency chain is *minimal* if the strict subterms of each t_i are terminating under $\rightarrow_{\mathcal{R}}$.

This definition corresponds to the first-order definition, except that a case for β -reduction is used, and that we explicitly require that $t_i \rightarrow_{in}^* s_{i+1}$: this is necessary because t_i may be an application rather than a functional term, and consequently may not be marked.

Theorem 5.7. *If \mathcal{R} is non-terminating, there is a minimal dependency chain over $\text{DP}(\mathcal{R})$.*

Proof. Given any non-terminating term, let q_{-1} be a minimal-sized subterm that is still non-terminating (q_{-1} is *MNT*, or *Minimal Non-Terminating*). We make the observations:

- (i) *If an MNT term is reduced at a non-top position, the result is either also MNT, or terminating.* This holds because, if $q = C[s] \rightarrow_{\mathcal{R}} C[t]$ because $s \rightarrow_{\mathcal{R}} t$, and t is non-terminating, then so is s , contradicting minimality of q unless $C = \square_\sigma$.
- (ii) *If $u \rightarrow_{in}^* v$, then $u^\sharp \rightarrow_{in}^* v^\sharp$.* This holds by the nature of an internal step.

For any $i \in \mathbb{N} \cup \{-1\}$, let q_i be a MNT term, and $t_i = q_i^\sharp$. Then q_i is not an abstraction, as abstractions can only be reduced by reducing their immediate subterm, contradicting minimality. For the same reason q_i cannot have the form $x \cdot u_1 \cdots u_n$ with x a variable, or $f(u_1, \dots, u_n) \cdot u_{n+1} \cdots u_m$ with f a constructor symbol. What remains are the forms:

- (A) $q_i = (\lambda x. u) \cdot v \cdot w_1 \cdots w_n$;
- (B) $q_i = f(v_1, \dots, v_n) \cdot v_{n+1} \cdots v_m$ with $f \in \mathcal{D}$.

We consider an infinite reduction starting in q_i . By minimality of q_i eventually a headmost step must be taken. In case (A) this must be a β -step because the left-hand sides of rules have the form $f(\vec{l}_1) \cdot \vec{l}_2$; therefore, the reduction has the form $q_i \rightarrow_{\mathcal{R}}^* (\lambda x. u') \cdot v' \cdot w'_1 \cdots w'_n \rightarrow_{\beta} u'[x := v'] \cdot w'_1 \cdots w'_n \rightarrow_{\mathcal{R}} \dots$. Since also $u[x := v] \cdot w_1 \cdots w_n \rightarrow_{\mathcal{R}}^* u'[x := v'] \cdot w'_1 \cdots w'_n$ the immediate beta-reduct of q_i is non-terminating as well. There are two sub-cases:

- If $n > 0$, this reduct is MNT by (i); in this case choose $q_{i+1} := u[x := v] \cdot w_1 \cdots w_n$ and let $\rho_{i+1}, s_{i+1}, t_{i+1} := \text{beta}, u_i, u_{i+1}$. Note that $s_{i+1}^\sharp = s_{i+1}$ and $t_{i+1}^\sharp = t_{i+1}$, and that case 3a of the definition of a dependency chain is satisfied.

- If $n = 0$, let w be a minimal-sized subterm of u where $w[x := v]$ is still non-terminating. By minimality of q_i both w and v are terminating, so $FV(w)$ contains x , but not $w = x$. Since w is not a variable, $(w[x := v])^\# = w^\#[x := v]$. By minimality of w , also $w[x := v]$ is MNT (its direct subterms have the form $w'[x := v]$ for a subterm w' of w). Case 3b is satisfied with $u_{i+1} := w[x := v]$ and $\rho_{i+1}, s_{i+1}, t_{i+1} := \mathbf{beta}, u_i, u_{i+1}^\#$.

Note that in both sub-cases, case 4 is also satisfied, since $t_i = s_{i+1}$.

In case (B), $q_i = f(v_1, \dots, v_n) \cdot v_{n+1} \cdots v_m$, we can always find a rule $l \rightarrow r$ and term $q'_i = l\gamma \cdot v'_{j+1} \cdots v'_m$ such that $q_i \rightarrow_{in}^* q'_i$, and $r\gamma \cdot v'_{j+1} \cdots v'_m$ is still non-terminating. Choose $s_{i+1} := q'_i$; requirement 4 from Definition 5.6 is satisfied by (ii). Since the rules were completed, we can assume that either $m = j$ or r is not an abstraction: if $r = \lambda x. r'$ and $m > j$ then $r\gamma \cdot v'_{j+1} \cdots v'_m$ is a β -redex, and (like above) may be reduced immediately without losing termination; the same result would have been obtained with the rule $l \cdot x \rightarrow r'$.

If $m > j$, then by (i) $r\gamma \cdot v'_{j+1} \cdots v'_m$ is MNT. Consequently, $\text{head}(r\gamma)$ cannot be a variable or a functional term $g(\vec{w})$ with g a constructor symbol: either $r\gamma$ is headed by an abstraction, or by a functional term with root symbol in \mathcal{D} . Since r itself is not an abstraction, its head must be a variable or a functional term with defined root symbol. Either way, $\rho_{i+1} := l \cdot x_{j+1} \cdots x_m \rightsquigarrow r \cdot x_{j+1} \cdots x_m$ is a dependency pair. Let $q_{i+1} := r\gamma \cdot v'_{j+1} \cdots v'_m$, and $t_{i+1} := q_{i+1}$ (which equals $q_{i+1}^\#$ as this is an application). Requirement 2 is satisfied.

Finally, if $m = j$, then $q'_i = l\gamma$ and $r\gamma$ is non-terminating. Let p be the smallest subterm of r such that $p[\vec{x} := \vec{c}]\gamma$ is non-terminating, where $\{\vec{x}\} = FV(p) \setminus FV(r)$. Then p is not a variable, for each $\gamma(x)$ is a subterm of $l\gamma$, and therefore terminating (and the c_σ do not reduce). Thus, the immediate subterms of $p[\vec{x} := \vec{c}]\gamma$ all have the form $p'[\vec{x} := \vec{c}]\gamma$ with $p \triangleright p'$, and are therefore terminating by minimality of p : $p[\vec{x} := \vec{c}]\gamma$ is MNT. As observed before, this can only be the case if this term is headed by an abstraction or by a functional term with a defined root symbol. And that can only be the case if p is either headed by a functional term with defined root symbol, or is an application headed by a variable which is free in r (as r has no subterms $(\lambda x. u) \cdot v$). Thus, $p[\vec{x} := \vec{c}]$ is a closed candidate term of r . As $p[\vec{x} := \vec{c}]\gamma$ is non-terminating, it is not a strict subterm of the MNT term $l\gamma$, so $\rho_{i+1} := l^\# \rightsquigarrow p[\vec{x} := \vec{c}]^\#$ is a dependency pair. Choose $q_{i+1} := p[\vec{x} := \vec{c}]\gamma$ and $t_{i+1} := q_{i+1}^\# = p[\vec{x} := \vec{c}]^\# \gamma$ (since $p[\vec{x} := \vec{c}]$ is not a variable). We see that in this case, too, requirement 2 is satisfied. \square

Example 5.8. As we will see, **twice** does not admit a dependency chain. As an example of a system which does admit one, consider the AFS with the following three rules:

$$\mathbf{f}(\mathbf{o}) \rightarrow \mathbf{g}(\lambda x. \mathbf{f}(x), \mathbf{a}) \quad \mathbf{g}(F, \mathbf{b}) \rightarrow F \cdot \mathbf{o} \quad \mathbf{a} \rightarrow \mathbf{b}$$

This system has four dependency pairs:

$$\mathbf{f}^\#(\mathbf{o}) \rightsquigarrow \mathbf{g}^\#(\lambda x. \mathbf{f}(x), \mathbf{a}) \quad \mathbf{f}^\#(\mathbf{o}) \rightsquigarrow \mathbf{f}^\#(\mathbf{c}_{\text{nat}}) \quad \mathbf{f}^\#(\mathbf{o}) \rightsquigarrow \mathbf{a}^\# \quad \mathbf{g}^\#(F, \mathbf{b}) \rightsquigarrow F \cdot \mathbf{o}$$

The rules admit an infinite reduction: $\mathbf{f}(\mathbf{o}) \rightarrow \mathbf{g}(\lambda x. \mathbf{f}(x), \mathbf{a}) \rightarrow \mathbf{g}(\lambda x. \mathbf{f}(x), \mathbf{b}) \rightarrow (\lambda x. \mathbf{f}(x)) \cdot \mathbf{o} \rightarrow_\beta \mathbf{f}(\mathbf{o}) \rightarrow \dots$; following the steps in the proof of Theorem 5.7 (starting with $\mathbf{f}(\mathbf{o})$) we obtain the following dependency chain:

$$\begin{aligned} & (\mathbf{f}^\#(\mathbf{o}) \rightsquigarrow \mathbf{g}^\#(\lambda x. \mathbf{f}(x), \mathbf{a}) \quad , \quad \mathbf{f}^\#(\mathbf{o}) \quad , \quad \mathbf{g}^\#(\lambda x. \mathbf{f}(x), \mathbf{a}) \quad), \\ & (\mathbf{g}^\#(F, \mathbf{b}) \rightsquigarrow F \cdot \mathbf{o} \quad , \quad \mathbf{g}^\#(\lambda x. \mathbf{f}(x), \mathbf{b}) \quad , \quad (\lambda x. \mathbf{f}(x), \mathbf{a}) \cdot \mathbf{o} \quad), \\ & (\mathbf{beta} \quad , \quad (\lambda x. \mathbf{f}(x), \mathbf{a}) \cdot \mathbf{o} \quad , \quad \mathbf{f}^\#(\mathbf{o}) \quad), \\ & (\mathbf{f}^\#(\mathbf{o}) \rightsquigarrow \mathbf{g}^\#(\lambda x. \mathbf{f}(x), \mathbf{a}) \quad , \quad \mathbf{f}^\#(\mathbf{o}) \quad , \quad \mathbf{g}^\#(\lambda x. \mathbf{f}(x), \mathbf{a}) \quad), \\ & \dots \end{aligned}$$

Note that between the first and second step, a \rightarrow_{in} step is done to reduce \mathbf{a} to \mathbf{b} . Also note that in the third triple we use case 3b from Definition 5.6, with $w = \mathbf{f}(x)$.

The converse of Theorem 5.7 does not hold. Consider for instance the AFS with symbols $\mathbf{A} : [\mathbf{nat} \times \mathbf{nat}] \Rightarrow \mathbf{nat}$ and $\mathbf{B} : [\mathbf{nat} \Rightarrow \mathbf{nat} \Rightarrow \mathbf{nat}] \Rightarrow \mathbf{nat}$, and a single rule: $\mathbf{A}(x, x) \rightarrow \mathbf{B}(\lambda y z. \mathbf{A}(y, z))$. This (terminating!) AFS has a dependency pair $\mathbf{A}^\#(x, x) \rightsquigarrow \mathbf{A}^\#(\mathbf{c}_{\mathbf{nat}}, \mathbf{c}_{\mathbf{nat}})$, which gives a dependency chain $\mathbf{A}^\#(\mathbf{c}_{\mathbf{nat}}, \mathbf{c}_{\mathbf{nat}}) \rightsquigarrow \mathbf{A}^\#(\mathbf{c}_{\mathbf{nat}}, \mathbf{c}_{\mathbf{nat}}) \rightsquigarrow \dots$

We could try solving this problem by slightly altering the definition of closed candidate terms: instead of substituting a variable $x : \sigma$ by a symbol \mathbf{c}_σ , we could have replaced it with a symbol \mathbf{c}_x , substituting all bound variables with different symbols. This choice was made in for example the first definition of dependency pairs for HRSs [35]. But even with this change, Theorem 5.7 does not give an equivalence. Consider for instance the AFS with the following rules:

$$\begin{aligned} \mathbf{f}(x, y, \mathbf{s}(z)) &\rightarrow \mathbf{g}(\mathbf{h}(x, y), \lambda u. \mathbf{f}(u, x, z)) \\ \mathbf{h}(x, x) &\rightarrow \mathbf{f}(x, \mathbf{s}(x), \mathbf{s}(\mathbf{s}(x))) \end{aligned}$$

This system has three dependency pairs:

$$\begin{aligned} \mathbf{f}^\#(x, y, \mathbf{s}(z)) &\rightsquigarrow \mathbf{h}^\#(x, y) \\ \mathbf{f}^\#(x, y, \mathbf{s}(z)) &\rightsquigarrow \mathbf{f}^\#(\mathbf{c}_u, x, z) \\ \mathbf{h}^\#(x, x) &\rightsquigarrow \mathbf{f}^\#(x, \mathbf{s}(x), \mathbf{s}(\mathbf{s}(x))) \end{aligned}$$

We get the following dependency chain: $\mathbf{f}^\#(\mathbf{c}_u, \mathbf{s}(\mathbf{c}_u), \mathbf{s}(\mathbf{s}(\mathbf{c}_u))) \rightsquigarrow \mathbf{f}^\#(\mathbf{c}_u, \mathbf{c}_u, \mathbf{s}(\mathbf{c}_u)) \rightsquigarrow \mathbf{h}^\#(\mathbf{c}_u, \mathbf{c}_u) \rightsquigarrow \mathbf{f}^\#(\mathbf{c}_u, \mathbf{s}(\mathbf{c}_u), \mathbf{s}(\mathbf{s}(\mathbf{c}_u))) \rightsquigarrow \dots$. However, the AFS is terminating, intuitively because the bound variable destroys matching possibilities with the non-left-linear rule.

For this reason, we have chosen to use the more elegant method with symbols \mathbf{c}_σ instead of the slightly more powerful, but also a fair bit more cumbersome, \mathbf{c}_x . The latter style is less pleasant because of α -conversion: for example, $\mathbf{Cand}(\mathbf{f}(\lambda x. \mathbf{g}(x)))$ should contain $\mathbf{g}(\mathbf{c}_y)$ for all variables y . Thus, to preserve correctness of definitions and proofs, we would have to jump through a few hoops. However, all results in this paper also go through with such a definition; this was for instance explored in the shorter version of this paper [24].

The crucial point of both examples above is the combination of bound variables and non-left-linear rules. However, for left-linear AFSs, no such counterexample exists. Intuitively, this holds because replacing variables by a symbol \mathbf{c}_σ that does not occur in any left-hand side does not affect applicability of any rule. Thus, a dependency chain effectively produces an infinite reduction $|s_i| \rightarrow_{\mathcal{R}} \cdot \supseteq |t_i| \rightarrow_{\mathcal{R}}^* |s_{i+1}|$ (where $|u|$ replaces any $f^\#$ in a term u by its unmarked counterpart), and this implies the existence of an infinite $\rightarrow_{\mathcal{R}}$ reduction.

Theorem 5.9. *A left-linear AFS \mathcal{R} is terminating if and only if it does not admit a (minimal) dependency chain.*

Proof. Theorem 5.7 gives one direction. For the other direction, assume a left-linear AFS \mathcal{R} and suppose we have an infinite dependency chain (minimal or not). We construct an infinite $\rightarrow_{\mathcal{R}} \cdot \supseteq$ sequence, following roughly the intuition above. We note:

- (1) If l is a linear term not containing any symbols \mathbf{c}_σ , and γ is a substitution whose domain contains only variables in l , and if $l\gamma = s[\vec{x} := \vec{c}]$ for some term s and set of variables $\{\vec{x}\}$, then there is a substitution δ such that $l\delta = s$ and $\gamma = \delta[\vec{x} := \vec{c}]$.
- (2) If $s[\vec{x} := \vec{c}] \rightarrow_{\mathcal{R}} t$, then there exists some t' such that $s \rightarrow_{\mathcal{R}} t'$ and $t'[\vec{x} := \vec{c}] = t$.
- (3) If $s[\vec{x} := \vec{c}] = C[t]$, then there are C' and t' such that $s = C'[t']$ and $t'[\vec{x} := \vec{c}] = t$.

(1) states that, if a linear term l (typically the left-hand side of a rule) matches a term s with some \mathbf{c} -symbols in it, it also matches s with those symbols replaced by variables.

This holds by induction on l , assuming linearity over $\text{dom}(\gamma)$: the cases where l is a variable are straightforward (if $l \in \text{dom}(\gamma)$ take $\delta = [l := s]$, otherwise let $\delta := \emptyset$), if l is an abstraction $\lambda x. l'$ the induction hypothesis suffices (x cannot occur in domain or range of δ , for then it would also hold for γ), and if l is an application or functional term we use the linearity. For example the functional case, if $l = f(l_1, \dots, l_n)$, then let γ_i be the restriction of γ to $FV(l_i)$ for $1 \leq i \leq n$; by the induction hypothesis we find suitable δ_i , and by linearity of l each of those l_i has different variables, so $\delta := \delta_1 \cup \dots \cup \delta_n$ is well-defined.

(2) states that, if a term with some variables replaced by \mathbf{c} -symbols reduces, then the original term reduces in a similar way. This holds by induction on the size of s . When the reduction is done in a subterm, the statement follows easily with the induction hypothesis (immediate subterms of $s[\vec{x} := \vec{c}]$ have the form $s'[\vec{x} := \vec{c}]$ with $s \triangleright s'$, so the induction hypothesis is applicable). In the base case, a β -step is easy, and if, for some rule $l \rightarrow r$ and substitution γ , the term $s[\vec{x} := \vec{c}] = l\gamma$, then by left-linearity of \mathcal{R} we may use (1): there is a substitution δ such that $s = l\delta \rightarrow_{\mathcal{R}} r\delta =: t'$; certainly $r\delta[\vec{x} := \vec{c}] = r\gamma = t$ as required.

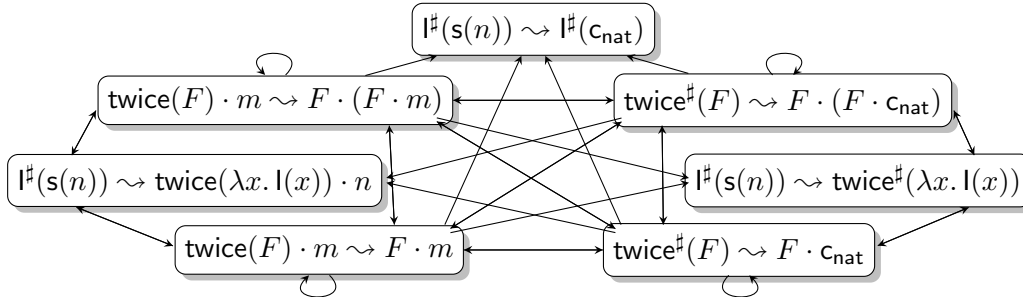
(3) follows by induction on the size of C : if C is the empty context take $t' := s$, otherwise use the induction hypothesis; for instance if $C[] = f(u_1, \dots, D_i[], \dots, u_n)$, then $s = f(s_1, \dots, s_i, \dots, s_n)$ (with each $s_j[\vec{x} := \vec{c}] = u_j$), and by the induction hypothesis on s_j there are D'_i, t' such that $C' := f(s_1, \dots, D'_i, \dots, s_n)$ and t' satisfy the requirement.

Now suppose there is a dependency chain $[(\rho_i, s_i, t_i) \mid i \in \mathbb{N}]$, and define $s'_0 := |s_0|$ (that is, s_0 with all marks removed). For all $i \in \mathbb{N}$, suppose $s'_i[\vec{x}_0 := \vec{c}, \dots, \vec{x}_{i-1} := \vec{c}] = |s_i|$. Whether ρ_i is **beta** or a dependency pair, $|s_i| \rightarrow_{\mathcal{R}} C_i[u_i]$ for some term u_i and context C_i , such that $|t_i| = u_i[\vec{x}_i := \vec{c}]$ for some variables \vec{x}_i . By (2), (3) also $s'_i \rightarrow_{\mathcal{R}} C'_i[u'_i]$ and $u'_i[\vec{x}_0 := \vec{c}, \dots, \vec{x}_i := \vec{c}] = |t_i|$. By (2) we can find s'_{i+1} such that $u'_i \rightarrow_{\mathcal{R}}^* s'_{i+1}$ and $s'_{i+1}[\vec{x}_0 := \vec{c}, \dots, \vec{x}_i := \vec{c}] = |s_{i+1}|$. Thus, s'_0 is non-terminating: $s'_0 \rightarrow_{\mathcal{R}}^+ C'_0[s'_1] \rightarrow_{\mathcal{R}}^+ C'_1[s'_2] \rightarrow_{\mathcal{R}}^+ \dots$ \square

5.3. The Dependency Graph. As in the first-order case, we use a dependency graph to organise the dependency pairs. The notions are very similar to the first-order definitions.

The *dependency graph* of an AFS \mathcal{R} is a graph with the dependency pairs of \mathcal{R} as nodes, and an edge from node $l \rightsquigarrow p$ to node $l' \rightsquigarrow p'$ if either $\text{head}(p)$ is a variable, or there are substitutions γ and δ such that $p\gamma \rightarrow_{\mathcal{R}, \text{in}}^* l'\delta$.

Example 5.10. The dependency graph of the AFS *twice*:



A *cycle* is a set \mathcal{C} of dependency pairs such that between every two pairs $\rho, \pi \in \mathcal{C}$ there is a non-empty path in the graph using only nodes in \mathcal{C} . A cycle that is not contained in

any other cycle is called a *strongly connected component* (SCC). To prove termination we must show that cycles in a dependency graph are “chain-free” (see Theorem 5.16). The requirement to add an edge from any node of the form $l \rightsquigarrow x \cdot r_1 \cdots r_n$ (with x a variable) to all other nodes is necessary by clause 3b in Definition 5.6: a dependency chain could have a dependency pair of the form $l \rightsquigarrow x \cdot \vec{r}$ followed by **beta**, and then any other dependency pair. Hence a rule with leading free variables in the right-hand side gives rise to many cycles.

A set of dependency pairs \mathcal{P} is called *chain-free* if there is no minimal dependency chain using only dependency pairs in $\mathcal{P} \cup \{\mathbf{beta}\}$.

Lemma 5.11. \emptyset is chain-free.

Proof. Given a dependency chain with all $\rho_i = \mathbf{beta}$, each $s_i \rightarrow_\beta C_i[s_{i+1}]$ for some context C_i , contradicting termination of the simply-typed λ -calculus. \square

Because the dependency graph cannot be computed in general, it is common to use *approximations* of the dependency graph, which have the same nodes but possibly more edges. A brute method to find an approximation is to have an edge between $l \rightsquigarrow p$ and $l' \rightsquigarrow p'$ if either the head of p is a variable, or if p and p' both have the form $f(s_1, \dots, s_n) \cdot s_{n+1} \cdots s_m$ for the same function symbol f . It is interesting to study more sophisticated methods to find approximations, but this is left for future work.

As stated in Section 3, we assume a finite set of rules, which leads to a finite set of dependency pairs. In Section 10.2 we will say a few words on extending the technique to systems with infinitely many rules (without having to deal with an infinite graph).

Lemma 5.12. Let G be an approximation of the dependency graph of an AFS \mathcal{R} . Suppose that every SCC in G is chain-free. Then \mathcal{R} is terminating.

Proof. Since DP is finite, any (minimal) dependency chain $[(\rho_i, s_i, t_i) \mid i \in \mathbb{N}]$ has at least one dependency pair ρ_i which occurs infinitely often. Note that if $n < m$ then there is a path in G from ρ_n to ρ_m (if $\rho_n, \rho_m \neq \mathbf{beta}$). Therefore, there is a path in G from ρ_i to itself, and hence ρ_i is on a cycle. Let \mathcal{C} be the SCC containing ρ_i . Then all ρ_j with $j > i$ and $\rho_j \neq \mathbf{beta}$ are in \mathcal{C} : from each such ρ_j there is a path to ρ_i and back. But then, $\{\rho_j \mid j \geq i\}$ is a minimal dependency chain in \mathcal{C} , so \mathcal{C} is not chain-free, contradicting the assumption! Thus, there is no dependency chain, and therefore by Theorem 5.7, \mathcal{R} is terminating. \square

Example 5.13. The dependency graph (approximation) of *twice* from Example 5.10 has only one SCC:

$$\left\{ \begin{array}{ll} l^\#(s(n)) \rightsquigarrow \text{twice}(\lambda x. l(x)) \cdot n & \text{twice}^\#(F) \rightsquigarrow F \cdot (F \cdot c_{\text{nat}}) \\ l^\#(s(n)) \rightsquigarrow \text{twice}^\#(\lambda x. l(x)) & \text{twice}^\#(F) \rightsquigarrow F \cdot c_{\text{nat}} \\ \text{twice}(F) \cdot m \rightsquigarrow F \cdot (F \cdot m) & \text{twice}(F) \cdot m \rightsquigarrow F \cdot m \end{array} \right\}$$

Therefore *twice* is terminating if this set, which we shall call $\mathcal{C}_{\text{twice}}$, is chain-free.

5.4. Reduction Triples. The challenge, then, is to prove that given sets of dependency pairs are chain-free. We use the following definition:

Definition 5.14. A *reduction triple* consists of a quasi-ordering \succeq , a sub-relation \succeq_1 of \succeq , and a well-founded ordering \succ , all defined on terms built over $\mathcal{F}_c^\#$, such that:

- (1) \succeq and \succ are *compatible*: either $\succ \cdot \succeq \subseteq \succ$ or $\succeq \cdot \succ \subseteq \succ$;
- (2) \succeq , \succeq_1 and \succ are all *stable* (closed under substitution);

- (3) \succeq_1 is *monotonic*: (if $s \succeq_1 t$ and s, t share a type, then $C[s] \succeq_1 C[t]$ for all $C[\]$);
- (4) \succeq_1 contains **beta** (always $(\lambda x. s) \cdot t \succeq_1 s[x := t]$).

A *reduction pair* is a pair (\succeq, \succ) such that $(\succeq, \succeq, \succ)$ is a reduction triple; this corresponds to the first-order notion of a reduction pair. The reduction triple is a generalisation of this notion, where \succeq itself is not required to be monotonic; we will need a non-monotonic \succeq in Section 5.5 to compare terms with different types. This notion of a reduction triple is similar to the one which appears in [15].

To deal with subterm reduction in dependency chains, an additional definition is needed.

Definition 5.15 (Limited Subterm Property). \succeq has the *limited subterm property* if the following requirement is satisfied: *for all variables x and terms s, t, u such that $s \succeq u \triangleright x$, there is a substitution γ such that $(\lambda x. s) \cdot t \succeq u^\sharp[x := t]\gamma$.*

Intuitively, the substitution γ can be used to replace free variables in u which are bound in s by the corresponding constants c_σ . However, we will also use a more liberal replacement of those variables, hence the general γ .

The following theorem shows how reduction triples are used with dependency pairs.

Theorem 5.16. *A set $\mathcal{P} = \mathcal{P}_1 \uplus \mathcal{P}_2$ of dependency pairs is chain-free if \mathcal{P}_2 is chain-free, and there is a reduction triple $(\succeq, \succeq_1, \succ)$ such that:*

- $l \succ p$ for all $l \rightsquigarrow p \in \mathcal{P}_1$,
- $l \succeq p$ for all $l \rightsquigarrow p \in \mathcal{P}_2$,
- $l \succeq_1 r$ for all $l \rightarrow r \in \mathcal{R}$,
- either \mathcal{P} is non-collapsing or \succeq has the limited subterm property.

Here, a set \mathcal{P} of dependency pairs is called *non-collapsing* if all elements of \mathcal{P} are non-collapsing. Symmetrically, \mathcal{P} is *collapsing* if it contains at least one collapsing pair $l^\sharp \rightarrow x \cdot \vec{p}$.

Proof. Towards a contradiction, suppose there is such a reduction triple $(\succeq, \succeq_1, \succ)$, but \mathcal{P} admits a minimal dependency chain; since \mathcal{P}_2 is chain-free, infinitely many ρ_i are in \mathcal{P}_1 .

If \mathcal{P} is non-collapsing, then the chain may start with some **beta** steps, but once some $\rho_i \in \text{DP}$, all ρ_j with $j > i$ must also be in DP, because the head of each t_i is a functional term, rather than an abstraction. Thus, for each j either $s_j \succeq t_j \succeq s_{j+1}$, or (if $\rho_j \in \mathcal{P}_1$) even $s_j \succ t_j \succeq s_{j+1}$, contradicting well-foundedness of \succ (the latter happens infinitely often).

Alternatively, suppose \mathcal{P} is collapsing, and \succeq has the limited subterm property. Let $[(\rho_i, s_i, t_i) \mid i \geq j]$ be a dependency chain over \mathcal{P} ; if $\rho_j \in \mathcal{P}_1$ then $s_j \succ t_j \succeq s_{j+1}$, if $\rho_j \in \mathcal{P}_2$ then $s_j \succeq t_j \succeq s_{j+1}$ and if $\rho_j = \text{beta}$ then (by the limited subterm property) there is a substitution δ such that $s_j \succeq t_j \delta \succeq s_{j+1} \delta$. Since $[(\rho_i, s_i \delta, t_i \delta) \mid i \geq j+1]$ is also a dependency chain we can continue this reasoning recursively. We obtain a decreasing \succeq sequence with infinitely many \succ steps, which contradicts well-foundedness of \succ . \square

Theorem 5.16 can be used to prove that every SCC in the dependency graph approximation of an AFS is chain-free; termination follows with Lemma 5.12. In Section 9 we will give an algorithm similar to the algorithm in Theorem 4.10.

Example 5.17. Termination of **twice** is proved if there is a reduction triple $(\succeq, \succeq_1, \succ)$ with the limited subterm property, such that $l \succeq_1 r$ for all rules, and $l \succ p$ for every dependency pair in $\mathcal{C}_{\text{twice}}$ from Example 5.13 (choosing $\mathcal{P}_2 = \emptyset$, which is chain-free).

For left-linear AFSs, where the existence of a minimal dependency chain characterises termination by Theorem 5.7, a terminating AFS always has a suitable reduction pair.

Theorem 5.18. *A left-linear AFS is terminating if and only if there is a reduction triple $(\succeq, \succeq_1, \succ)$ such that $l \succ p$ for every $l \rightsquigarrow p \in \text{DP}$, and $l \succeq_1 r$ for every $l \rightarrow r \in \mathcal{R}$, and \succeq has the limited subterm property.*

Proof. By Lemma 5.12 and Theorem 5.16, termination of \mathcal{R} follows if such a reduction triple exists. For the other direction, let $s \succeq t$ if $|s| \rightarrow_{\mathcal{R}}^* |t|$, and let $s \succ t$ if $|s| (\rightarrow_{\mathcal{R}} \cdot \sqsupseteq)^+ |t|$, where \sqsupseteq is the (reflexive) subterm relation where bound variables which become free are replaced with symbols c_σ , and $|u|$ removes marks from u . It is evident that $(\succeq, \succeq_1, \succ)$ is a reduction triple, that $l \succ p$ for all dependency pairs $l \rightsquigarrow p$ and $l \succeq_1 r$ for all rules $l \rightarrow r$. Moreover, \succeq has the limited subterm property with γ the substitution $[\bar{y} := \bar{c}]$. \square

5.5. Type Changing. The situation so far is not completely satisfactory, because both \succeq and \succ may have to compare terms of different types. Consider for example the dependency pair $\text{twice}^\sharp(F) \rightsquigarrow F \cdot c_{\text{nat}}$, where the left-hand side has a functional type and the right-hand side does not. Moreover, the comparison in the definition of limited subterm property may concern terms of different types. This is problematic because term orderings do not usually relate terms of arbitrary different types; neither any version of the higher-order path ordering [19, 4] nor monotonic algebras [32] are equipped to do this.

A solution is to manipulate the ordering requirements. Let (\succeq, \succ) be a reduction pair (so a pair such that $(\succeq, \succeq_1, \succ)$ is a reduction triple). Define \geq, \geq_1 and $>$ as follows:

- $s > t$ if there are fresh variables x_1, \dots, x_n and terms u_1, \dots, u_m such that $s \cdot x_1 \cdots x_n \succ t \cdot u_1 \cdots u_m$ and both sides have some base type;
- $s \geq t$ if there are fresh variables x_1, \dots, x_n and terms u_1, \dots, u_m such that $s \cdot x_1 \cdots x_n R t \cdot u_1 \cdots u_m$ and both sides have some base type; here R is the union of $\succeq, \succ \cdot \succeq$ and $\succeq \cdot \succ$;
- $s \geq_1 t$ if $s \geq t$ and s and t have the same type.

Lemma 5.19. *$(\geq, \geq_1, >)$ as generated from a reduction pair (\succeq, \succ) is a reduction triple.*

Proof. We make the following observations:

- (1) if $s \geq_1 t$ then by monotonicity $s\vec{x} \succeq t\vec{x}$;
- (2) if $s > t$ then for any \vec{u} there are \vec{v} such that $s \cdot \vec{u} \succ t \cdot \vec{v}$ (by stability of \succ);
- (3) if $s \geq t$ then for any \vec{u} there are \vec{v} such that either $s \cdot \vec{u} \succeq t \cdot \vec{v}$ or $s \cdot \vec{u} \succ \cdot \succeq t \cdot \vec{v}$ or $s \cdot \vec{u} \succeq \cdot \succ t \cdot \vec{v}$ (by stability of both \succ and \succeq).

Each of the required properties on \geq, \geq_1 , and $>$ now follows easily from the properties on \succeq and \succ . For example transitivity of $>$: if $s > t > u$, then there are terms \vec{v} such that $s \cdot \vec{x} \succ t \cdot \vec{v}$, and by (2) there are terms \vec{w} such that $t \cdot \vec{v} \succ u \cdot \vec{w}$; by transitivity of \succ , therefore, $s \cdot \vec{x} \succ u \cdot \vec{w}$, so $s > u$. Well-foundedness of $>$ follows from (2) and well-foundedness of \succ . For stability, note that if $s > t$ and γ is a substitution, then for fresh variables \vec{x} (which do not occur in domain or range of γ) also $s \cdot \vec{x} \succ t \cdot \vec{v}$, so $(s\gamma) \cdot \vec{x} = (s \cdot \vec{x})\gamma \succ (t \cdot \vec{v})\gamma = (t\gamma) \cdot (\vec{v}\gamma)$; stability of \geq is similar. \geq_1 is included in \geq by (1), and contains **beta** because \succeq does. For compatibility, and for transitivity of \geq , we use a case distinction on which form of \geq is used, and transitivity of both \succ and \succeq , as well as compatibility between the two. \square

The relations \geq and $>$ are not necessarily computable, but they do not need to be: we will only use specific instances. To prove some set of dependency pairs \mathcal{P} chain-free, we can choose for every pair $l \rightsquigarrow p \in \mathcal{P}$ a corresponding base-type pair $\bar{l} \rightsquigarrow \bar{p}$, and prove either $\bar{l} \succeq \bar{p}$ or $\bar{l} \succ \bar{p}$. For example, we could assign $\bar{l} := l \cdot x_1 \cdots x_n$ and $\bar{p} := p \cdot c_{\sigma_1} \cdots c_{\sigma_m}$. This is the choice we will use in examples in this paper. Other choices for \bar{p} are also possible.

We assume a systematic way of choosing $\bar{l} \rightsquigarrow \bar{p}$ given $l \rightsquigarrow p$.

To make sure that \geq has the limited subterm property, we consider a base-type version of subterm reduction, which has a strong relation with β -reduction.

Definition 5.20. $\triangleright^!$ is the relation on base-type terms (and $\triangleright^!$ its reflexive closure) generated by the following clauses:

- $(\lambda x. s) \cdot t_0 \cdots t_n \triangleright^! u$ if $s[x := t_0] \cdot t_1 \cdots t_n \triangleright^! u$;
- $f(s_1, \dots, s_m) \cdot t_1 \cdots t_n \triangleright^! u$ if $s_i \cdot \vec{c} \triangleright^! u$ for some i ;
- $s \cdot t_1 \cdots t_n \triangleright^! u$ if $t_i \cdot \vec{c} \triangleright^! u$ for some i (s may have any form).

Here, $s \cdot \vec{c}$ is a term s applied to constants c_σ of the right types. We say (\succeq, \succ) respects $\triangleright^!$ if $\triangleright^!$ is contained in $(\succeq \cup \succ)^*$. Note that, since \succeq contains **beta**, the first clause is not likely to give problems. $\triangleright^!$ is interesting because if $s \triangleright t$ and s has base type, then there are terms u_1, \dots, u_n and a substitution γ on domain $FV(t) \setminus FV(s)$ such that $s \triangleright^! t\gamma \cdot u_1 \cdots u_n$ (this is easy to see with induction on the size of s). Consequently, \geq has the limited subterm property if (\succeq, \succ) respects $\triangleright^!$ and $f(\vec{x}) \succeq f^\sharp(\vec{x})$ for all $f \in \mathcal{D}$ (the *marking property*).

Using Theorem 5.16 and the reduction triple generated from a reduction pair, we obtain:

Theorem 5.21. A set of dependency pairs $\mathcal{P} = \mathcal{P}_1 \uplus \mathcal{P}_2$ is chain-free if \mathcal{P}_2 is chain-free and there is a reduction pair (\succeq, \succ) such that:

- (1) $\bar{l} \succ \bar{p}$ for all $l \rightsquigarrow p \in \mathcal{P}_1$;
- (2) $\bar{l} \succeq \bar{p}$ for all $l \rightsquigarrow p \in \mathcal{P}_2$;
- (3) $l \succeq r$ for all $l \rightarrow r \in \mathcal{R}$;
- (4) if \mathcal{P} is collapsing, then (\succeq, \succ) respects $\triangleright^!$, and $f(\vec{x}) \succeq f^\sharp(\vec{x})$ for all $f \in \mathcal{D}$.

Note that the theorem does not use the generated (and possibly not computable) triple directly; we prove $\bar{l} \succ \bar{p}$ or $\bar{l} \succeq \bar{p}$ for a specific choice of \bar{l} and \bar{p} . The generated triple is merely used in the reasoning that justifies Theorem 5.21.

Example 5.22. To prove that $\mathcal{C}_{\text{twice}}$ is chain-free it suffices to find a reduction pair (\succeq, \succ) such that $l \succeq r$ for all rules, (\succeq, \succ) respects $\triangleright^!$ and satisfies the marking property, and:

$$\begin{array}{lll} l^\sharp(s(n)) & \succ & \text{twice}(\lambda x. l(x)) \cdot n & \text{twice}^\sharp(F) \cdot x & \succ & F \cdot (F \cdot c_{\text{nat}}) \\ l^\sharp(s(n)) & \succ & \text{twice}^\sharp(\lambda x. l(x)) \cdot c_{\text{nat}} & \text{twice}^\sharp(F) \cdot x & \succ & F \cdot c_{\text{nat}} \\ \text{twice}(F) \cdot m & \succ & F \cdot (F \cdot m) & \text{twice}(F) \cdot m & \succ & F \cdot m \end{array}$$

This completes the basis of dynamic dependency pairs for AFSs.

At this point, we might ask: *what have we gained?* Is it easier to use Theorem 5.21 than to use a conventional approach like CPO [4]? Can we even find a reduction pair which respects $\triangleright^!$? And if so, couldn't we use the same reduction pair without dependency pairs?

The answer to these questions will be explored in the coming sections. First (Section 6), we will consider an extension limited to *fully extended, left-linear AFSs*. With this restriction, we can weaken the limited subterm property, and obtain a variation of usable rules. Next, in Section 7, we will study two ways to find a suitable reduction pair: using interpretations in

a *weakly monotonic algebra*, and *argument functions*, a generalisation of argument filterings. Finally, in Section 8 we will see additional ways to prove chain-freeness of a set \mathcal{P} if \mathcal{P} is *non-collapsing*. All results are combined in the algorithm of Section 9.

6. DEPENDENCY PAIRS FOR LOCAL AFSs

In this section we consider local AFSs, and define formative rules for local AFSs. We add tags to symbols below a λ , and prove that we only need the limited subterm property for tagged symbols. We use this to weaken the requirements on a reduction pair.

The limited subterm property is weaker than the requirements used in Theorem 5.21: subterm reduction *only* has to be done following β -reduction. That is, we only need it for terms which occur below a λ -abstraction, when a bound variable is substituted.

To exploit this property, we will pay special attention to *local* AFSs. In a local AFS we can (mostly) avoid reducing terms below an abstraction. Knowing this, the limited subterm property only requires that $f(s_1, \dots, s_n) \cdot s_{n+1} \cdots s_m \succeq s_i \cdot \vec{c}$ for symbols f which cannot, at that time, be reduced anyway. This makes it possible to use for instance argument filterings, as we will see in this section and Section 7.2. As a bonus, locality also allows us to define *formative rules*, a variation of usable rules. These are discussed in Section 6.4.

6.1. Intuition. The idea to tag symbols and rules in order to weaken the limited subterm property originates in the notion of *weak reductions*, defined in [6] (following a definition from Howard in 1968). A weak reduction in the λ -calculus does not use steps between a λ -abstraction and its binder. This notion generalises to AFSs in the obvious way.

Consider AFSs where the left-hand sides of all rules are *linear* (so no free variables occur more than once), and *free of abstractions* (so the λ symbol does not occur in them). This limitation is not as strong as it might seem at first; the β -reduction “rule” is not included in this. As it turns out, we can prove the following statement:

Claim: *in a left-linear, and left-abstraction-free AFS, if there is a minimal dependency chain, then there is one where the reduction $t_i \rightarrow_{\mathcal{R}}^* s_{i+1}$ always uses only weak steps.*

To see why this matters, let us consider a colouring of the function symbols. In a given term s , make all symbol occurrences either red or green: red if the symbol occurs between an abstraction and its binder, green otherwise. So if $s = C[f(t_1, \dots, t_n)]$, make the f red if some t_i contains freely a variable which is bound in s , green if not. We say s is well-coloured if it uses this colouring. Colour the rules in the same way; by the restrictions, the left-hand sides are entirely green, while the right-hand sides may contain red symbols.

Now consider a weak reduction step on a well-coloured term. If the term is reduced by a coloured rule, then the result is also well-coloured. If the term is reduced with a β -step, then the result may have some red symbols outside an abstraction; however, it can become well-coloured again by painting these red symbols green. We never have to paint green symbols red. Inventing notation, we can summarise this as follows:

Claim: *if $s \rightarrow_{\mathcal{R}, \text{weak}} t$, then $\text{colour}(s) \rightarrow_{\mathcal{R}_{\text{colour}}}^* \rightarrow_{\text{make_green}}^* \text{colour}(t)$.*

Combining the two claims, we can colour dependency chains. In the **beta**-with-subterm step (3b), which led to the need for the limited subterm property, we take a term which was originally below an abstraction, reduce it to a subterm which still contains the bound variable, and substitute it. Importantly, the subterm clause $u \succeq v$ can be derived with steps $\lambda x. s \triangleright s$, $s_1 \cdot s_2 \triangleright s_i$ and $f(s_1, \dots, s_n) \triangleright s_i$, where the f is always a red symbol.

Considering the red and green symbols as different symbols altogether (related only by the *make_green* rules) we thus see that it will not give problems to use an argument filtering, provided we use it only for the green symbols!

This summarises the ideas which we shall use to simplify the limited subterm property. Since colours do not work well in papers, we will use tags: a red symbol f corresponds with a symbol f^- , and a green symbol remains unchanged. Moreover, if we focus on the colours, and forget about the weak reductions, it turns out that we do not need to require that the left-hand sides of rules contain no λ -abstractions at all: it suffices if the rules are *local*.

6.2. Local AFSs. Both to weaken the limited subterm property, and for formative rules, we shall restrict attention to so-called local AFSs where, intuitively, matching is purely local. This means that to apply a rule we do not have to check whether two subterms are equal, or whether a symbol occurs in a subterm. Locality combines the restrictions that the system is *left-linear* and *fully extended*. A left-linear, left-abstraction-free AFS is always local.

The locality restriction appears in the literature both for HRSs [31, 5], where a pattern is called local if it is fully extended and linear, and for combinatory reduction systems (CRSs) in [29]; the latter definition is slightly different but has a similar underlying intuition. The definition for AFSs here follows [31, 5], although the definition of full-extendedness is technically (but not conceptually) different from the one for HRSs. We will use locality to be able to (mostly) postpone reductions below an abstraction. In the explanations below, we will argue that left-linearity and full-extendedness are both necessary to do this.

Left-linearity When a system is not left-linear, a reduction deep inside a term may be needed to create a topmost redex. For instance, consider the non-left-linear AFS with rules $\{f(x, x) \rightarrow b, a \rightarrow b\}$. In the reduction $f(\lambda x. a, \lambda x. b) \rightarrow f(\lambda x. b, \lambda x. b) \rightarrow b$ a reduction below an abstraction is necessary to create the syntactic equality required for the f -rule. Thus, this step cannot be postponed.

Full-Extendedness We say a term l is *fully extended* if free variables in l do not occur below an abstraction; a rule $l \rightarrow r$ is called fully extended if l is.

For the intuition of this restriction, consider a rule $f(\lambda x. y) \rightarrow y$. This rule does not match a term $f(\lambda x. s(x))$, since y cannot be instantiated with $s(x)$, as x is bound. Nor does $f(\lambda x. F \cdot x)$ match this term, since $s(x)$ does not instantiate the application $F \cdot x$. Whenever the left-hand side of a rule contains a free variable below an abstraction, this variable matches only subterms which do not contain the abstraction-variable. Therefore, such a rule could require a reduction deep inside a term to create a topmost redex. For example, in an AFS with rules $\{f(\lambda x. y) \rightarrow y, g(x, y) \rightarrow a\}$, we cannot postpone the first step in the reduction $f(\lambda x. h(\lambda y. g(x, y))) \rightarrow f(\lambda x. h(\lambda y. a)) \rightarrow h(\lambda y. a)$, as it is needed to create the second redex.

This notion of fully extended mostly corresponds with the definition for HRSs; there, however, a rule $f(\lambda x. F \cdot x) \rightarrow r$ *does* match $f(\lambda x. s(x))$, so such rules are also accepted.

Definition 6.1. An AFS $(\mathcal{F}, \mathcal{R})$ is *local* if all $l \rightarrow r \in \mathcal{R}$ are left-linear and fully extended.

Example 6.2. Our running example, $\mathcal{R}_{\text{twice}}$, is local, since all left-hand sides of rules are linear and fully extended (in fact, they contain no abstractions at all).

To demonstrate the prominence of local AFSs, in the 2011 version of the Termination Problem Data Base, used in the annual termination competition [38], 138 out of 156 benchmarks in the higher-order category are local (in fact, these are all left-abstraction-free).

6.3. Tagging Unreducible Symbols. Obviously, when there are rules where the left-hand side contains an abstraction, such as $f(\lambda x. g(x), F) \rightarrow r$, it may be impossible to avoid reducing inside an abstraction in order to create a redex. However, the colouring intuition still goes through; we merely need to “paint symbols green” a few times more.

Following the colouring intuition, we will mark all function symbols which occur between a λ -abstraction and its binder with a special tag (“colouring red”). The symbol can only be reduced by removing the tag first (“painting green”).

Definition 6.3. Let \mathcal{F}^- be the set $\{f^- : \sigma \mid f : \sigma \in \mathcal{F}\}$, so a set containing a “tagged” symbol f^- for all function symbols $f \in \mathcal{F}$. For a set of variables Z , define tag_Z as follows:

$$\begin{aligned} \text{tag}_Z(x) &= x \\ \text{tag}_Z(c_\sigma) &= c_\sigma \\ \text{tag}_Z(s \cdot t) &= \text{tag}_Z(s) \cdot \text{tag}_Z(t) \\ \text{tag}_Z(\lambda x. s) &= \lambda x. \text{tag}_{Z \cup \{x\}}(s) \\ \text{tag}_Z(f(s_1, \dots, s_n)) &= \begin{cases} f(\text{tag}_Z(s_1), \dots, \text{tag}_Z(s_n)) & \text{if } FV(f(\vec{s})) \cap Z = \emptyset \\ f^-(\text{tag}_Z(s_1), \dots, \text{tag}_Z(s_n)) & \text{if } FV(f(\vec{s})) \cap Z \neq \emptyset \end{cases} \end{aligned}$$

We denote $\text{tag}(s) := \text{tag}_\emptyset(s)$. Define $\mathcal{R}^{\text{tag}} := \{l \rightarrow \text{tag}_\emptyset(r) \mid l \rightarrow r \in \mathcal{R}\} \cup \{f^-(x_1, \dots, x_n) \rightarrow f(x_1, \dots, x_n) \mid f^- \in \mathcal{F}^-\}$.

Note that, apart from the untagging rules, \mathcal{R}^{tag} isn’t all that different from \mathcal{R} : $\text{tag}(r)$ is almost exactly r , only the symbols below an abstraction may be marked with a $-$ sign.

Example 6.4. $\text{tag}(f(\lambda x. g(x, g(o)))) = f(\lambda x. g^-(x, g(o)))$.

Example 6.5. Consider our running example $\mathcal{R}_{\text{twice}}$ (with completed rules):

$$\begin{array}{ll} l(o) \rightarrow o & \text{twice}(F) \rightarrow \lambda y. F \cdot (F \cdot y) \\ l(s(n)) \rightarrow s(\text{twice}(\lambda x. l(x)) \cdot n) & \text{twice}(F) \cdot m \rightarrow F \cdot (F \cdot m) \end{array}$$

We have seen that $\mathcal{R}_{\text{twice}}$ is local. \mathcal{R}^{tag} consists of the following rules:

$$\begin{array}{ll} l(o) \rightarrow o & \text{twice}(F) \rightarrow \lambda y. F \cdot (F \cdot y) \\ l(s(n)) \rightarrow s(\text{twice}(\lambda x. l^-(x)) \cdot n) & \text{twice}(F) \cdot n \rightarrow F \cdot (F \cdot n) \\ o^- \rightarrow o & s^-(n) \rightarrow s(n) \\ l^-(n) \rightarrow l(n) & \text{twice}^-(F) \rightarrow \text{twice}(F) \end{array}$$

That is, the rules from \mathcal{R} , with a tag added to the l symbol which occurs below an abstraction, and furthermore the untagging rules. In termination proofs we can typically ignore the rules $f^-(\vec{x}) \rightarrow f(\vec{x})$ where f^- does not occur in the right-hand side of any rule (in this example: $o^- \rightarrow o$, $s^-(n) \rightarrow s(n)$ and $\text{twice}^-(F, n) \rightarrow \text{twice}(F, n)$), as they have little function.

In the proofs later on in this section, we will use the following properties of \mathcal{R}^{tag} :

Lemma 6.6. $\text{tag}_{X \cup Y}(s) \rightarrow_{\mathcal{R}^{\text{tag}}}^* \text{tag}_X(s)$ for any set of rules \mathcal{R} ; if the variables in Y don’t occur in s even $\text{tag}_{X \cup Y}(s) = \text{tag}_X(s)$.

Proof. Easy induction on the size of s ; we only use the untagging rules $f^-(\vec{x}) \rightarrow f(\vec{x})$. \square

Lemma 6.7. $\text{tag}(s)\gamma^{\text{tag}} = \text{tag}(s\gamma)$ where $\gamma^{\text{tag}} = [x := \text{tag}(\gamma(x)) \mid x \in \text{dom}(\gamma)]$.

Proof. We prove by induction on the size of s : for any set of variables Z , whose members do not occur in either domain or range of γ , we have $\text{tag}_Z(s)\gamma^{\text{tag}} = \text{tag}_Z(s\gamma)$.

If s is a variable not in $\text{dom}(\gamma)$, both sides are just s .

If s is a variable in $\text{dom}(\gamma)$, we must see that $\text{tag}(\gamma(s)) = \text{tag}_Z(\gamma(s))$, which holds by the second part of Lemma 6.6.

If s is an application $t \cdot u$, then $\text{tag}_Z(s)\gamma^{\text{tag}} = (\text{tag}_Z(t)\gamma^{\text{tag}}) \cdot (\text{tag}_Z(u)\gamma^{\text{tag}})$, which by the induction hypothesis equals $\text{tag}(t\gamma) \cdot \text{tag}(u\gamma) = \text{tag}((t\gamma) \cdot (u\gamma)) = \text{tag}(s\gamma)$.

If $s = f(s_1, \dots, s_n)$ the induction hypothesis on each of the s_i also suffices because $Z \cap FV(s) = Z \cap FV(s\gamma)$, which is easy to see by the requirements on Z .

Finally, if $s = \lambda y. s'$, then $\text{tag}_Z(s)\gamma^{\text{tag}} = (\lambda y. \text{tag}_{Z \cup \{y\}}(s'))\gamma^{\text{tag}} = \lambda y. (\text{tag}_{Z \cup \{y\}}(s')\gamma^{\text{tag}})$, which by the induction hypothesis equals $\lambda y. \text{tag}_{Z \cup \{y\}}(s'\gamma) = \text{tag}_Z(s\gamma)$ as required. \square

Lemma 6.8. If Z is a set of variables, s, t terms and x a variable not in Z or $FV(t)$, then $\text{tag}_{Z \cup \{x\}}(s)[x := \text{tag}_Z(t)] \rightarrow_{\mathcal{R}^{\text{tag}}}^* \text{tag}_Z(s[x := t])$ for any set of rules \mathcal{R} .

Proof. By induction on the size of s .

If $s = x$, then both sides are equal to $\text{tag}_Z(t)$.

If s is another variable y , then both sides are just y .

If $s = u \cdot v$ we use the induction hypothesis: $\text{tag}_{Z \cup \{x\}}(s)[x := \text{tag}_Z(t)] = \text{tag}_{Z \cup \{x\}}(u)[x := \text{tag}_Z(t)] \cdot \text{tag}_{Z \cup \{x\}}(v)[x := \text{tag}_Z(t)] \rightarrow_{\mathcal{R}^{\text{tag}}}^* \text{tag}_Z(u[x := t]) \cdot \text{tag}_Z(v[x := t]) = \text{tag}_Z(u[x := t] \cdot v[x := t]) = \text{tag}_Z(s[x := t])$.

If $s = \lambda y. u$, we use the second part of Lemma 6.6: $\text{tag}_{Z \cup \{x\}}(s)[x := \text{tag}_Z(t)] = \lambda y. \text{tag}_{Z \cup \{x, y\}}(u)[x := \text{tag}_Z(t)]$, which by Lemma 6.6 equals $\lambda y. \text{tag}_{Z \cup \{x, y\}}(u)[x := \text{tag}_{Z \cup \{y\}}(t)] \rightarrow_{\mathcal{R}^{\text{tag}}}^* \lambda y. \text{tag}_{Z \cup \{y\}}(u[x := t]) = \text{tag}_Z(s[x := t])$ by the induction hypothesis.

Finally, if $s = f(u_1, \dots, u_n)$, then there is little to do if x does not occur in s : by the second part of Lemma 6.6, $\text{tag}_{Z \cup \{x\}}(s)[x := \text{tag}_Z(t)] = \text{tag}_Z(s)[x := \text{tag}_Z(t)]$ and since x does not occur in either s or $\text{tag}_Z(s)$, this is exactly $\text{tag}_Z(s[x := t])$. So assume that $x \in FV(s)$; then $\text{tag}_{Z \cup \{x\}}(s)[x := \text{tag}_Z(t)] = f^-(\text{tag}_{Z \cup \{x\}}(u_1), \dots, \text{tag}_{Z \cup \{x\}}(u_n))[x := \text{tag}_Z(t)]$, which by the induction hypothesis reduces to $f^-(\text{tag}_Z(u_1[x := t]), \dots, \text{tag}_Z(u_n[x := t]))$. If variables of Z occur in $s[x := t]$ this is exactly $\text{tag}_Z(s[x := t])$, otherwise it reduces in one step to $f(\text{tag}_Z(u_1[x := t]), \dots, \text{tag}_Z(u_n[x := t])) = \text{tag}_Z(s)$. \square

The following lemma expresses that a reduction to a term of a certain form l can be done by only reducing subterms headed by untagged (“green”) symbols. Later on, we will use this to see that the reduction $t_i \rightarrow_{\mathcal{R}}^* s_{i+1}$ in a dependency chain can be assumed to reduce only untagged symbols.

Lemma 6.9. Let \mathcal{R} be a local AFS, l a linear, fully extended term and γ a substitution on domain $FV(l)$. If s is terminating and $s \rightarrow_{\mathcal{R}}^* l\gamma$, then there is a substitution δ such that $\text{tag}(s) \rightarrow_{\mathcal{R}^{\text{tag}}}^* l\delta^{\text{tag}}$, where $\delta^{\text{tag}} = [x := \text{tag}(\delta(x)) \mid x \in \text{dom}(\delta)]$, and $\delta(x) \rightarrow_{\mathcal{R}}^* \gamma(x)$ for all x .

Proof. Towards an induction hypothesis, we will prove the lemma for a term l which is linear and fully extended in the variables in $\text{dom}(\gamma)$, and such that $\text{dom}(\gamma) \subseteq FV(l)$; l may have more variables which do not occur in this domain, and which it is not necessarily linear and fully extended in. We use induction first on s , using $\rightarrow_{\mathcal{R}} \cup \triangleright$ (this is well-founded because s is terminating by assumption), second on the length of the reduction $s \rightarrow_{\mathcal{R}}^* l\gamma$.

First suppose l is a variable in $\text{dom}(\gamma)$, so $\gamma = [l := \gamma(l)]$. Choose $\delta(l) = s$. Then certainly $\delta(l) \rightarrow_{\mathcal{R}}^* \gamma(l)$, and $\text{tag}(s) \rightarrow_{\mathcal{R}^{\text{tag}}}^* \text{tag}(s) = l\delta^{\text{tag}}$. If l is a variable not in $\text{dom}(\gamma)$, and $s \rightarrow_{\mathcal{R}}^* l\gamma = l$ without headmost steps, then γ is empty and $s = l$; indeed $\text{tag}(s) = l \rightarrow_{\mathcal{R}^{\text{tag}}}^* l$.

Next, let l be an abstraction $\lambda x. l'$ and suppose $s = \lambda x. s'$ and $s' \rightarrow_{\mathcal{R}}^* l'\gamma$. Since l is fully extended in the variables of $\text{dom}(\gamma)$, this l' contains no variables in $\text{dom}(\gamma)$; that is, γ is empty, and we must see that $\text{tag}(s) \rightarrow_{\mathcal{R}}^* l$. By the induction hypothesis $\text{tag}(s') \rightarrow_{\mathcal{R}}^* l'$, and therefore indeed $\text{tag}(s) = \lambda x. \text{tag}_{\{x\}}(s') \rightarrow_{\mathcal{R}^{\text{tag}}}^* \lambda x. \text{tag}(s')$ by Lemma 6.6, $\rightarrow_{\mathcal{R}^{\text{tag}}}^* \lambda x. l' = l$.

If $l = f(l_1, \dots, l_n)$ and $s = f(s_1, \dots, s_n)$ and each $s_i \rightarrow_{\mathcal{R}}^* l_i\gamma$, then by linearity of l each l_i has different variables (at least, insofar as $\text{dom}(\gamma)$ is concerned). We can write $\gamma = \gamma_1 \cup \dots \cup \gamma_n$ with each γ_i the restriction of γ to $\text{dom}(l_i)$; all γ_i have disjoint domains. Then also $s_i \rightarrow_{\mathcal{R}}^* l_i\gamma_i$, so by the induction hypothesis there are $\delta_1, \dots, \delta_n$ such that each $\text{tag}(s_i) \rightarrow_{\mathcal{R}^{\text{tag}}}^* l_i\delta_i^{\text{tag}}$, and always $\delta_i(x) \rightarrow_{\mathcal{R}}^* \gamma_i(x)$. The induction step holds with $\delta := \delta_1 \cup \dots \cup \delta_n$.

If $l = l_1 \cdot l_2$ and $s = s_1 \cdot s_2$ and each $s_i \rightarrow_{\mathcal{R}}^* l_i\gamma$, we use linearity in the same way.

If none of these cases hold, the reduction $s \rightarrow_{\mathcal{R}}^* l\gamma$ must use a headmost step, so $s \rightarrow_{\mathcal{R}}^* u \rightarrow_{\mathcal{R}} v \rightarrow_{\mathcal{R}}^* l\gamma$, and either:

- $u = (\lambda x. w) \cdot q_0 \cdots q_n$ and $v = w[x := q_0] \cdot q_1 \cdots q_n$ ($n \geq 0$), or
- $u = l'\gamma' \cdot q_1 \cdots q_n$ and $v = r'\gamma' \cdot q_1 \cdots q_n$ for some $l' \rightarrow r', \gamma, q_1, \dots, q_n$ ($n \geq 0$)

We can safely assume that the reduction $s \rightarrow_{\mathcal{R}}^* u$ does not use any headmost steps.

In the first case, s must have the form $(\lambda x. w') \cdot q'_0 \cdots q'_n$ with $w' \rightarrow_{\mathcal{R}}^* w$ and each $q'_i \rightarrow_{\mathcal{R}}^* q_i$. But then also $w'[x := q'_0] \cdot q'_1 \cdots q'_n \rightarrow_{\mathcal{R}}^* v \rightarrow_{\mathcal{R}}^* l\gamma$; by the first induction hypothesis we find a suitable δ such that $s \rightarrow_{\beta} w'[x := q'_0] \cdot q'_1 \cdots q'_n \rightarrow_{\mathcal{R}^{\text{tag}}}^* l\delta^{\text{tag}}$.

In the second case, let $l'' := l \cdot x_1 \cdots x_n$ for fresh variables x_1, \dots, x_n , and let $\gamma'' := \gamma \cup [x_1 := q_1, \dots, x_n := q_n]$. Since $l' \rightarrow r'$ is a rule, l'' is both linear and fully extended, and the reduction to $l''\gamma'' = u$ is shorter than the original reduction; by the second induction hypothesis we find χ such that $\text{tag}(s) \rightarrow_{\mathcal{R}^{\text{tag}}}^* l''\chi^{\text{tag}} = l'\chi^{\text{tag}} \cdot (x_1\chi^{\text{tag}}) \cdots (x_n\chi^{\text{tag}})$, where each $\chi(y)$ reduces to $\gamma'(y)$ for $y \in \text{dom}(\gamma')$, and $\chi(x_i) \rightarrow_{\mathcal{R}}^* q_i$.

Now, $l''\chi^{\text{tag}} \rightarrow_{\mathcal{R}^{\text{tag}}}^* \text{tag}(r')\chi^{\text{tag}} \cdot (\vec{x}\chi^{\text{tag}})$, which by Lemma 6.7 = $\text{tag}(r'\chi) \cdot (\vec{x}\chi^{\text{tag}}) = \text{tag}((r' \cdot \vec{x})\chi)$. By simply removing all tags, every $\rightarrow_{\mathcal{R}^{\text{tag}}}^*$ step can be translated to a $\rightarrow_{\mathcal{R}}^*$ step on untagged terms, and therefore we also see that $s \rightarrow_{\mathcal{R}}^* l''\chi \rightarrow_{\mathcal{R}} (r \cdot \vec{x})\chi$, and by the choice of χ we know: $(r \cdot \vec{x})\chi \rightarrow_{\mathcal{R}}^* (r'\gamma') \cdot \vec{q} \rightarrow_{\mathcal{R}}^* l\gamma$. Therefore we can apply the first induction hypothesis, and see that $\text{tag}(s) \rightarrow_{\mathcal{R}^{\text{tag}}}^+ \text{tag}((r \cdot \vec{x})\chi) \rightarrow_{\mathcal{R}^{\text{tag}}}^* l\delta^{\text{tag}}$, for a suitable δ . \square

In Section 6.5, Lemma 6.9 will play an essential role in the construction of a “tagged dependency chain”. But first, let us consider *formative rules*, another gain from locality.

6.4. Formative Rules. Recall that in the first-order setting it is not required to prove $l \succeq r$ for *all* rewrite rules: to prove that a set of dependency pairs \mathcal{P} is chain-free it suffices to consider only its *usable rules*. The definition of usable rules cannot easily be extended to our setting, because we normally have to deal with collapsing dependency pairs. Therefore we take a different approach with the same goal of restricting attention to rules which are in some way relevant to a set of dependency pairs. Where usable rules are defined from the right-hand sides of dependency pairs, our *formative rules* are based on the left-hand sides.

The intuition behind formative rules is that (due to left-linearity and full-extendedness), only the formative rules of some rule $l \rightarrow r$ can contribute to the creation of its pattern.

We consider a fixed set of rules \mathcal{R} , which has already been completed. The formative rules are a subset of \mathcal{R}^+ , which is the set $\mathcal{R} \cup \{l \cdot x_1 \cdots x_n \rightarrow r \cdot x_1 \cdots x_n \mid l \rightarrow r \in \mathcal{R}, \text{ all } x_i \text{ fresh variables, } r \text{ not an abstraction and } l \cdot x_1 \cdots x_n \text{ well-typed}\}$.

Definition 6.10 (Formative Rules). Let X be a set of variables, and s a β -normal term (that is, s has no subterms $(\lambda x. t) \cdot u$ such that for any subterm $x \cdot t$ of s with $x \in \mathcal{V}$, either x is not free in s , or $x \in X$). Let $Symb_X(s)$ be recursively defined as follows:

$$\begin{aligned} Symb_X(\lambda y. s : \sigma) &= \{\langle ABS, \sigma \rangle\} \cup Symb_{X \cup \{y\}}(s) \\ Symb_X(f(s_1, \dots, s_n) \cdot s_{n+1} \cdots s_m : \sigma) &= \{\langle f, \sigma \rangle\} \cup Symb_X(s_1) \cup \dots \cup Symb_X(s_m) \\ Symb_X(y \cdot s_1 \cdots s_n : \sigma) &= \{\langle VAR, \sigma \rangle \cup Symb_X(s_1) \cup \dots \cup Symb_X(s_n) \mid y \in X, n \geq 0\} \\ Symb_X(y) &= \emptyset \quad (y \in \mathcal{V} \setminus X) \end{aligned}$$

Note that in a local AFS, all left-hand sides of the rules satisfy these constraints for $X = \emptyset$.

For $a \in \mathcal{F} \cup \{ABS, VAR\}$, we say a term $s : \sigma$ has form $\langle a, \sigma \rangle$ if either $a = ABS$ and s is an abstraction, or $a \in \mathcal{F}$ and s can be written $a(\vec{t}) \cdot \vec{u}$, or $s = x \cdot \vec{t}$ for some variable x (and a may be anything). A pair $\langle a, \sigma \rangle$ with $a \in \mathcal{F} \cup \{ABS, VAR\}$ is called a *typed symbol*.

For two typed symbols A, B , write $A \sqsubseteq_{fo} B$ if there is a rule $l \rightarrow r \in \mathcal{R}^+$ such that r has form A , and $B \in Symb_\emptyset(l)$. Let \sqsubseteq_{fo}^* denote the reflexive-transitive closure of \sqsubseteq_{fo} .

The *formative symbols* of a term s are those typed symbols B such that $A \sqsubseteq_{fo}^* B$ for some $A \in Symb_\emptyset(s)$ (if defined).

The *formative rules* of a term s , notation $FR(s)$, are those rules $l \rightarrow r \in \mathcal{R}^+$ such that r has form B for some formative symbol B of s .

The set of formative rules of a dependency pair, $FR(f(l_1, \dots, l_n) \cdot l_{n+1} \cdots l_m \rightsquigarrow p)$, is defined as $\bigcup_{1 \leq i \leq m} FR(l_i)$. For a set \mathcal{P} of dependency pairs, $FR(\mathcal{P}) = \bigcup_{l \rightsquigarrow p \in \mathcal{P}} FR(l \rightsquigarrow p)$.

Note that in a finite system it is easy to calculate the formative symbols of a term, and consequently the formative rules can be found automatically.

Example 6.11. Recall the rules for the (completed) system twice:

$$\begin{array}{ll} (A) \quad l(o) \rightarrow o & (C) \quad \text{twice}(F) \rightarrow \lambda y. F \cdot (F \cdot y) \\ (B) \quad l(s(n)) \rightarrow s(\text{twice}(\lambda x. l(x)) \cdot n) & (D) \quad \text{twice}(F) \cdot m \rightarrow F \cdot (F \cdot m) \end{array}$$

Here $\mathcal{R}^+ = \mathcal{R}$. In this context, let $l = s(n)$. Then $Symb_\emptyset(l) = \{\langle s, \text{nat} \rangle\}$, and:

- (B) and (D) both have form $\langle s, \text{nat} \rangle$, so $\langle s, \text{nat} \rangle \sqsubseteq_{fo} \langle s, \text{nat} \rangle, \langle l, \text{nat} \rangle, \langle \text{twice}, \text{nat} \rangle$
- (D) also has forms $\langle l, \text{nat} \rangle$ and $\langle \text{twice}, \text{nat} \rangle$, but no other rules do

Thus, the formative symbols of l are exactly $\langle s, \text{nat} \rangle$, $\langle l, \text{nat} \rangle$ and $\langle \text{twice}, \text{nat} \rangle$. (B) and (D), but not (A) and (C), are formative rules of l . Observing that a dependency pair with left-hand side $\text{twice}(F) \cdot n$ or $\text{twice}^\sharp(F)$ has no formative rules (since $Symb_\emptyset(F) = Symb_\emptyset(n) = \emptyset$), the formative rules of the SCC $\mathcal{C}_{\text{twice}}$ from Example 5.13 are (B) and (D).

Example 6.12. For an example that uses multiple types, and more rules of functional type, consider the system with symbols

$$\begin{array}{ll} \text{cons} : [(\text{nat} \Rightarrow \text{nat}) \times \text{funlist}] \Rightarrow \text{funlist} & \text{nil} : \text{funlist} \\ \text{head} : [\text{funlist}] \Rightarrow \text{nat} \Rightarrow \text{nat} & \text{tail} : [\text{funlist}] \Rightarrow \text{funlist} \\ \text{true} : \text{bool} & \text{false} : \text{bool} \\ \text{test} : [\text{nat} \Rightarrow \text{nat}] \Rightarrow \text{bool} & s : [\text{nat}] \Rightarrow \text{nat} \\ \text{if} : [\text{bool} \times (\text{nat} \Rightarrow \text{string}) \times (\text{nat} \Rightarrow \text{string})] \Rightarrow \text{nat} \Rightarrow \text{string} \end{array}$$

And rules:

$$\begin{array}{ll} (A) \quad \text{if}(\text{true}, F_1, F_2) \rightarrow F_1 & (D) \quad \text{head}(\text{cons}(F, t)) \rightarrow F \\ (B) \quad \text{if}(\text{false}, F_1, F_2) \rightarrow F_2 & (E) \quad \text{tail}(\text{cons}(F, t)) \rightarrow t \\ (C) \quad \text{test}(\lambda x. s(x)) \rightarrow \text{true} & \end{array}$$

For \mathcal{R}^+ , we add the rules:

$$\begin{aligned} (F) \quad & \text{if}(\text{true}, F_1, F_2) \cdot x \rightarrow F_1 \cdot x \\ (G) \quad & \text{if}(\text{false}, F_1, F_2) \cdot x \rightarrow F_2 \cdot x \\ (H) \quad & \text{head}(\text{cons}(F, t)) \cdot x \rightarrow F \cdot x \end{aligned}$$

This is a contrived example, to demonstrate all aspects of formative rules in one system. We consider the formative rules of the dependency pair $\text{if}(\text{true}, F_1, F_2) \cdot x \rightsquigarrow F_1 \cdot x$. That is, $FR(\text{true})$, since the free variables F_1, F_2, x do not have formative rules. We observe:

- $Symb_\emptyset(\text{true}) = \langle \text{true}, \text{bool} \rangle$
- rule (C) is the only rule with form $\langle \text{true}, \text{bool} \rangle$, so $\langle \text{true}, \text{bool} \rangle \sqsubseteq_{fo} \langle \text{test}, \text{bool} \rangle$, $\langle ABS, \text{nat} \Rightarrow \text{nat} \rangle$, $\langle s, \text{nat} \rangle$, $\langle VAR, \text{nat} \rangle$, that is, the elements of $Symb_\emptyset(\text{test}(\lambda x. s(x)))$;
- $\langle ABS, \text{nat} \Rightarrow \text{nat} \rangle \sqsubseteq_{fo} \langle \text{head}, \text{nat} \Rightarrow \text{nat} \rangle$, $\langle \text{cons}, \text{funlist} \rangle$ by rule (D), and rule (H) has both form $\langle s, \text{nat} \rangle$ and $\langle VAR, \text{nat} \rangle$, so these two $\sqsubseteq_{fo} \langle \text{head}, \text{nat} \rangle$, $\langle \text{cons}, \text{funlist} \rangle$; no other rule has a form $\langle \text{test}, \text{bool} \rangle$, $\langle ABS, \text{nat} \Rightarrow \text{nat} \rangle$, $\langle s, \text{nat} \rangle$ or $\langle VAR, \text{nat} \rangle$
- (D) also has form $\langle \text{head}, \text{nat} \Rightarrow \text{nat} \rangle$ (but we already know that all symbols in the left-hand side are formative symbols of true), and only (E) has form $\langle \text{cons}, \text{funlist} \rangle$, so the latter $\sqsubseteq_{fo} \langle \text{tail}, \text{funlist} \rangle$, $\langle \text{cons}, \text{funlist} \rangle$
- Hence, the formative symbols of the given dependency pair are:
 $\langle \text{true}, \text{bool} \rangle$, $\langle \text{test}, \text{bool} \rangle$, $\langle s, \text{nat} \rangle$, $\langle VAR, \text{nat} \rangle$, $\langle \text{head}, \text{nat} \rangle$, $\langle \text{cons}, \text{funlist} \rangle$, $\langle \text{tail}, \text{funlist} \rangle$, $\langle ABS, \text{nat} \Rightarrow \text{nat} \rangle$, $\langle \text{head}, \text{nat} \Rightarrow \text{nat} \rangle$.

The formative rules are therefore (C), (D), (E) and (H).

For formative rules we have a result very similar to Lemma 6.9, both in nature and in proof.

Lemma 6.13. *Suppose \mathcal{R} is local, and let l be a β -normal, linear, fully extended term, which does not have leading free variables. Let γ be a substitution with domain $FV(l)$, and s a term which is terminating over $\rightarrow_{\mathcal{R}}$, and suppose $s \rightarrow_{\mathcal{R}}^* l\gamma$. Then there is a substitution δ on $FV(l)$ such that $s \rightarrow_{FR(l)}^* l\delta$, and moreover each $\delta(x) \rightarrow_{\mathcal{R}}^* \gamma(x)$.*

Proof. We will prove something slightly stronger, which implies the lemma. Let X be a set of variables, and l a β -normal term, linear in $FV(l) \setminus X$, and such that if a free variable x occurs inside an abstraction, or at the head of an application in l , then $x \in X$. Let γ be a substitution with domain $FV(l) \setminus X$, and s a terminating term such that $s \rightarrow_{\mathcal{R}}^* l\gamma$.

Let $FS_X(l)$ denote the set of typed symbols B such that $A \sqsubseteq_{fo}^* B$ for some $A \in Symb_X(l)$, and $FR_X(l)$ is the set of rules $l' \rightarrow r'$ in \mathcal{R}^+ such that r' has form B for some $B \in FS_X(l)$. We will find a substitution δ on $FV(l) \setminus X$ such that $s \rightarrow_{FR(l)}^* l\delta$, and always $\delta(x) \rightarrow_{\mathcal{R}}^* \gamma(x)$.

It is clear that, for $X = \emptyset$, the definitions of $FR_X(l)$ and $FR(l)$ coincide. Thus, the case $X = \emptyset$ implies the lemma – but for the induction step we will need a larger X .

Before proving this claim, let us make the following observations:

- (1) if $Symb_X(s) \subseteq Symb_Y(t)$, then $FS_X(s) \subseteq FS_Y(t)$, so $FR_X(s) \subseteq FR_Y(t)$
- (2) if the variables of Y do not occur in s , then $Symb_{X \cup Y}(s) = Symb_X(s)$
- (3) if $s \triangleright t$ and $Y = FV(t) \setminus FV(s)$, then $Symb_{X \cup Y}(t) \subseteq Symb_X(s)$
- (4) if $u \rightarrow v \in FR_X(s)$, then $FS_\emptyset(u) \subseteq FS_X(s)$

All of these are obvious by considering the respective definitions.

Now we have all the preparations to prove the required result, using induction on s with $\rightarrow_{\mathcal{R}} \cup \triangleright$. Because the rules have been completed and l is β -normal, we can first transform the reduction $s \rightarrow_{\mathcal{R}}^* l\gamma$ into a reduction which never takes a headmost step with a rule $l' \rightarrow \lambda x. r'$ which is not also a topmost step (we can replace these steps one by one, and by

induction on s with $\rightarrow_{\mathcal{R}}$ we eventually obtain a reduction without such steps). Having done this, we use a second induction, on the length of the reduction $s \rightarrow_{\mathcal{R}}^* l\gamma$. Now, we can prove the claim. Consider the form of l .

If l is a variable in $\text{dom}(\gamma)$, then $\gamma = [l := \gamma(l)]$; choosing $\delta := [l := s]$ we are done.

If $l = x \cdot l_1 \cdots l_n$ with $x \in X$, and $s = x \cdot s_1 \cdots s_n$ and each $s_i \rightarrow_{\mathcal{R}}^* l_i \gamma$, then by linearity of l over $\text{dom}(\gamma)$ we can write $\gamma = \gamma_1 \cup \dots \cup \gamma_n$ where γ_i is the restriction of γ to $FV(l_i)$. By the induction hypothesis we can find $\delta_1, \dots, \delta_n$ such that each $s_i \rightarrow_{FR_X(l_i)}^* l_i \delta_i$ and $\delta_i \rightarrow_{\mathcal{R}}^* \gamma_i$. Choose $\delta := \delta_1 \cup \dots \cup \delta_n$; this is well-defined because all δ_i have disjoint domains. By (3) and (1) each $FR_X(l_i) \subseteq FR_X(l)$, so indeed $s \rightarrow_{FR_X(l)}^* l\delta$, and also each $\delta(x) \rightarrow_{\mathcal{R}}^* \gamma(x)$.

If $l = f(l_1, \dots, l_m) \cdot l_{m+1} \cdots l_n$, and $s = f(s_1, \dots, s_m) \cdot s_{m+1} \cdots s_n$ and each $s_i \rightarrow_{\mathcal{R}}^* l_i \gamma$, we use linearity in almost exactly the same way.

If $l = \lambda x. l'$ and $s = \lambda x. s'$ and $s' \rightarrow_{\mathcal{R}}^* l' \gamma$, then by assumption the term l' contains only variables in X , so $\text{dom}(\gamma) = \emptyset$; we must show that $s' \rightarrow_{FR_X(l)}^* l'$. By (3), (2) and (1), it suffices if $s' \rightarrow_{FR_X \cup \{x\}(l')}^* l'$, and this is exactly what the induction hypothesis gives us!

By the restrictions on l , it must have one of the forms above; if we are not yet done, therefore, the reduction $s \rightarrow_{\mathcal{R}}^* l\gamma$ uses a headmost step.

If s has the form $(\lambda x. t) \cdot u \cdot \vec{v}$, then the first headmost step must be a β -step: $s \rightarrow_{\mathcal{R}}^* (\lambda x. t') \cdot u' \cdot \vec{v} \rightarrow_{\beta} t'[x := u'] \cdot \vec{v}' \rightarrow_{\mathcal{R}}^* l\gamma$; we might as well β -reduce immediately, and have $s \rightarrow_{\beta} t[x := u] \cdot \vec{v} \rightarrow_{\mathcal{R}}^* t'[x := u'] \cdot \vec{v}' \rightarrow_{\mathcal{R}}^* l\gamma$; the first induction hypothesis gives a suitable δ .

If s does not have this form, there is at least one headmost step which is not a β -reduction. The reduction has a form $s \rightarrow_{\mathcal{R}}^* t \rightarrow_{\mathcal{R}} u \rightarrow_{\mathcal{R}}^* l\gamma$, where $t = l'\gamma' \cdot v_1 \cdots v_n \rightarrow_{\mathcal{R}} r'\gamma' \cdot v_1 \cdots v_n$ for some rule $l' \rightarrow r'$, substitution γ' and terms v_1, \dots, v_n ($n \geq 0$); we can choose t, u in such a way that the reduction $u \rightarrow_{\mathcal{R}}^* l\gamma$ does not contain any headmost steps other than perhaps β -steps. Let $l'' := l' \cdot x_1 \cdots x_n$ and $r'' := r' \cdot x_1 \cdots x_n$ for suitably typed fresh variables x_1, \dots, x_n ; then $l'' \rightarrow r''$ is in \mathcal{R}^+ , because we have made sure that either r' is not an abstraction, or $n = 0$. Let $\gamma'' := \gamma' \cup [x_1 := v_1, \dots, x_n := v_n]$. Then $t = l''\gamma''$ and $u = r''\gamma''$. Applying the second induction hypothesis on the reduction $s \rightarrow_{\mathcal{R}}^* l''\gamma''$, we find some substitution χ such that $s \rightarrow_{FR_0(l'')}^* l''\chi \rightarrow_{\mathcal{R}} r''\chi \rightarrow_{\mathcal{R}}^* r''\gamma'' \rightarrow_{\mathcal{R}}^* l\gamma$. Note that $FR_0(l'') \subseteq \mathcal{R}^+$, and that $\rightarrow_{\mathcal{R}^+}$ defines the same relation as $\rightarrow_{\mathcal{R}}$. Thus, $s \rightarrow_{\mathcal{R}}^* l''\chi \rightarrow_{\mathcal{R}} r''\chi$; we can apply the first induction hypothesis to find a suitable δ such that $r''\chi \rightarrow_{FR_X(l)}^* l\delta$.

Suppose $l'' \rightarrow r'' \in FR_X(l)$. Then by (4) and (1), also $FR_0(l'') \subseteq FR_X(l)$, so we have a reduction $s \rightarrow_{FR_X(l)}^* l''\chi \rightarrow_{FR_X(l)} r''\chi \rightarrow_{FR_X(l)}^* l\delta$, and we are done. To see that this is indeed the case, first suppose that $\text{head}(r')$ is a variable. Whatever the form of l is (since $l \notin \text{dom}(\gamma)$), $\text{Symb}_X(l)$ contains a pair $\langle f, \sigma \rangle$, where σ is the type of l (and also the type of s , l'' and r''), and $f \in \mathcal{F} \cup \{ABS, VAR\}$. We immediately see that $l'' \rightarrow r'' \in FR_X(l)$. Alternatively, if $\text{head}(r')$ is a function symbol, then u is not a β -redex; as the reduction $u \rightarrow_{\mathcal{R}}^* l\gamma$ does not use other headmost steps, we have $u \rightarrow_{\mathcal{R},in}^* l\gamma$, and $l = f(l_1, \dots, l_k) \cdot l_{k+1} \cdots l_m$, where f is also the head-symbol of r' . But then $\langle f, \sigma \rangle \in \text{Symb}_X(l)$, so also $l'' \rightarrow r'' \in FR_X(l)$. \square

Of course, Lemma 6.13 and Lemma 6.9 can be combined: the latter doesn't care *which* rules it is given, so if $s \rightarrow_{\mathcal{R}}^* l\gamma$, then there are δ, χ such that $s \rightarrow_{FR(l)}^* l\delta$, and $\text{tag}(s) \rightarrow_{FR(l)\text{tag}}^* l\chi^{\text{tag}}$ and each $\chi(x) \rightarrow_{FR(l)}^* \delta(x) \rightarrow_{\mathcal{R}}^* \gamma(x)$. In the following, we will use this combination of lemmas to see that, for local AFSs, a dependency chain can be assumed to use tagged steps and formative rules in the reduction $t_i \rightarrow_{\mathcal{R},in}^* s_{i+1}$.

Comment: The formative rules technique is also applicable to first-order rewriting, in particular for many-sorted TRSs (or for innermost rewriting where types may be added by [7]). However, we have not yet investigated whether the technique leads to an improvement in current state-of-the-art termination provers.

6.5. Revised Dependency Pair Results for Local AFSs. We may now revise the results from Section 5 to take locality into account. As before, we assume that the rules in \mathcal{R} are all completed, and let DP be the dependency pairs of \mathcal{R} . Because of Lemmas 6.9 and 6.13 we can consider an alternative definition of dependency chain.

Definition 6.14. A *tagged dependency chain* is a sequence $[(\rho_i, s_i, t_i) \mid i \in \mathbb{N}]$ with for all i :

- (1) $\rho_i \in \text{DP} \cup \{\text{beta}\}$
- (2) if $\rho_i = l_i \rightsquigarrow p_i \in \text{DP}$ then $s_i = l_i \gamma^{\text{tag}}$ and $t_i = \text{tag}(p_i) \gamma^{\text{tag}}$ for some substitution γ
- (3) if $\rho_i = \text{beta}$ then $s_i = \text{tag}((\lambda x. u) \cdot v \cdot w_1 \cdots w_k)$ and either
 - (a) $k > 0$ and $t_i = \text{tag}(u[x := v] \cdot w_1 \cdots w_k)$, or
 - (b) $k = 0$ and there exists a term w such that $u \geq w$ and $x \in FV(w)$ and $t_i = \text{tag}(w^\sharp[x := v])$, but $w \neq x$
- (4) $t_i \rightarrow_{FR(l_{i+1})^{\text{tag}}, in}^* s_{i+1}$

A tagged dependency chain is *minimal* if $\text{untag}(u)$ is terminating under $\rightarrow_{\mathcal{R}}$ for all strict subterms u of each t_i (where $\text{untag}(\)$ removes the $-$ tags).

This definition is similar to the original definition of a dependency chain, but uses tags for s_i and t_i and limits the rules in the \rightarrow_{in}^* reduction to the formative rules of the pattern which is created. We obtain the following variation of Theorem 5.9:

Theorem 6.15. A local AFS \mathcal{R} is non-terminating if and only if it admits a minimal tagged dependency chain.

Proof. If we remove the tags from a tagged dependency chain, we obtain a normal dependency chain. Since local AFSs are left-linear, Theorem 5.9 provides one direction.

For the other direction, we follow the proof of Theorem 5.7; in each step i we have a minimal non-terminating, untagged term q_i , and $t_i = \text{tag}(q_i^\sharp)$. If $\text{head}(q_i)$ is an abstraction we follow the proof of Theorem 5.7 to find q_{i+1} ; the requirements of Definition 6.14 are satisfied for $s_{i+1} := \text{tag}(q_i)$ and $t_{i+1} := \text{tag}(q_{i+1}^\sharp)$. Otherwise, let $q_i = f(v_1, \dots, v_m) \cdot v_{m+1} \cdots v_n$.

Since q_i is MNT, an infinite $\rightarrow_{\mathcal{R}}^*$ -reduction starting in q_i must eventually take a headmost step, say $q_i \rightarrow_{\mathcal{R}, in}^* l \gamma \cdot v'_{k+1} \cdots v'_n$ (with $k \geq m$), where $l \rightarrow r \in \mathcal{R}$ and $r \gamma \cdot v'_{k+1} \cdots v'_n$ is still non-terminating. Write $l = f(\tilde{l}_1, \dots, \tilde{l}_m) \cdot \tilde{l}_{m+1} \cdots \tilde{l}_k$; by left-linearity all \tilde{l}_j have disjunct free variables. Applying Lemmas 6.9 and 6.13 on each of the v_j and \tilde{l}_j (with the suitable part γ_j of γ), such a redex can be reached with tagged steps and formative rules: $\text{tag}(v_j) \rightarrow_{FR(\tilde{l}_j)^{\text{tag}}}^* \tilde{l}_j \delta_j^{\text{tag}}$, and $\delta_j \rightarrow_{\mathcal{R}}^* \gamma_j$. Choosing $\delta := \delta_1 \cup \dots \cup \delta_n$ we have that $r \delta \cdot v_{k+1} \cdots v_n \rightarrow_{\mathcal{R}}^* r \gamma \cdot v'_{k+1} \cdots v'_n$ is still non-terminating. Let $q'_i := l \delta \cdot v_{k+1} \cdots v_m$ and $s_{i+1} := l \delta^{\text{tag}} \cdot \text{tag}(v_{k+1}) \cdots \text{tag}(v_m)$ and continue the proof as before; in the resulting dependency pair $l_{i+1} \rightsquigarrow p_{i+1}$ all \tilde{l}_j are immediate subterms of l_{i+1} , so $FR(\tilde{l}_j) \subseteq FR(l_{i+1} \rightsquigarrow p_{i+1})$. We have $q_{i+1} = |p_{i+1}| \chi$ for some substitution χ , so $t_{i+1} := \text{tag}(q_{i+1}) = \text{tag}(p_{i+1}) \chi^{\text{tag}}$ as required, by Lemma 6.7. \square

Example 6.16. Consider once more the non-terminating system from Example 5.8

$$f(o) \rightarrow g(\lambda x. f(x), a) \quad g(F, b) \rightarrow F \cdot o \quad a \rightarrow b$$

Noting that \mathcal{R}^{tag} consists of the rules

$$\begin{array}{ll} f(o) \rightarrow g(\lambda x. f^-(x), a) & g(F, b) \rightarrow F \cdot o \\ a \rightarrow b & f^-(x) \rightarrow f(x) \end{array}$$

as well as some other rules $h^-(\vec{x}) \rightarrow h(\vec{x})$, we have the following tagged dependency chain:

$$\begin{array}{l} (f^\#(o) \rightsquigarrow g^\#(\lambda x. f^-(x), a), \quad f^\#(o), \quad g^\#(\lambda x. f^-(x), a)) \\ (g^\#(F, b) \rightsquigarrow F \cdot o, \quad g(\lambda x. f^-(x), b), \quad (\lambda x. f^-(x)) \cdot o) \\ (\text{beta}, \quad (\lambda x. f^-(x)) \cdot o, \quad f^\#(o)) \\ (f^\#(o) \rightsquigarrow g^\#(\lambda x. f^-(x), a), \quad f^\#(o), \quad g^\#(\lambda x. f^-(x), a)) \\ \dots \end{array}$$

Here, the **beta** step uses case 3b with $w = f(x)$.

As in Section 5 we consider the dependency graph of \mathcal{R} . A set $\mathcal{P} \subseteq \text{DP}$ is *tagged-chain-free* if there is no minimal tagged dependency chain using only dependency pairs from \mathcal{P} , and **beta**. As before, \emptyset is tagged-chain-free, and \mathcal{R} is terminating if and only if every SCC in a graph approximation is tagged-chain-free. Thus, we can use reduction triples which orient the parts of a *tagged* dependency chain. Importantly, this affects the limited subterm property.

Definition 6.17 (Tagged Subterm Property). \succeq has the *tagged subterm property* if the following requirement is satisfied: *for all variables x and terms s, t, u such that $s \succeq u \triangleright x$, there is a substitution γ such that $\text{tag}((\lambda x. s) \cdot t) \succeq \text{tag}(u^\# [x := t] \gamma)$.*

As we will see shortly, the tagged subterm property is an improvement over the limited subterm property because we do not have to take the subterms of untagged functional terms $f(\vec{s})$. It is easy to adapt the proof of Theorem 5.16 to obtain the following result:

Theorem 6.18. *A set $\mathcal{P} = \mathcal{P}_1 \uplus \mathcal{P}_2$ of dependency pairs is tagged-chain-free if \mathcal{P}_2 is tagged-chain-free and there is a reduction triple $(\succeq, \succeq_1, \succ)$ such that:*

- $l \succ \text{tag}(p)$ for all $l \rightsquigarrow p \in \mathcal{P}_1$;
- $l \succeq \text{tag}(p)$ for all $l \rightsquigarrow p \in \mathcal{P}_2$;
- $l \succeq_1 \text{tag}(r)$ for all $l \rightarrow r \in \text{FR}(\mathcal{P})$;
- $f^-(\vec{x}) \succeq_1 f(\vec{x})$ for all $f^- \in \mathcal{F}^-$;
- either \mathcal{P} is non-collapsing or \succeq has the tagged subterm property.

Proof. If the properties above are satisfied, then every minimal tagged dependency chain over \mathcal{P} leads to an infinite decreasing \succ -chain, contradicting well-foundedness of \succ . The elements of this proof are straightforward, following the proof of Theorem 5.16, except perhaps for the proof that there is a substitution δ such that $s_i \succeq t_i \delta$ when $s_i = \text{tag}((\lambda x. u) \cdot v \cdot \vec{w})$ and either $t_i = \text{tag}(q^\# [x := v])$ (if $|\vec{w}| = 0$ and $u \triangleright q \triangleright x$), or $t_i = \text{tag}(u[x := v] \cdot \vec{w})$.

The first case of this holds by the tagged subterm property: $\text{tag}(s_i) = \text{tag}((\lambda x. u) \cdot v) \succeq \text{tag}(q^\# [x := v] \gamma)$ for some γ , and this equals $\text{tag}(q^\# [x := v]) \gamma^{\text{tag}}$ by Lemma 6.7; let $\delta := \gamma^{\text{tag}}$.

For the second case, $\text{tag}(s_i) = (\lambda x. \text{tag}_{\{x\}}(u)) \cdot \text{tag}(v) \cdot \text{tag}(\vec{w}) \succeq \text{tag}_{\{x\}}(u)[x := \text{tag}(v)] \cdot \text{tag}(\vec{w})$ (since \succeq_1 includes **beta**), $\succeq \text{tag}(u[x := v]) \cdot \text{tag}(\vec{w}) = \text{tag}(t_i)$ by Lemma 6.8 and because always $f^-(\vec{x}) \succeq_1 f(\vec{x})$. \square

Theorem 5.18 also has a counterpart: if a local AFS is non-terminating, then there is a reduction triple which satisfies the constraints of Theorem 6.18 for $\mathcal{P} = \mathcal{P}_1 = \text{DP}$: if $(\succeq, \succeq_1, \succ)$ is the reduction triple from Theorem 5.18, let $s R' t$ if $\text{untag}(s) R \text{untag}(t)$. Then $(\succeq', \succeq'_1, \succ')$ satisfies the required properties.

Theorem 6.18 is comparable to Theorem 5.16, and as before, the result is likely not immediately usable due to typing problems. Moreover, it is not obvious that the tagged subterm property is really weaker than the limited subterm property. So to complete the work, we re-examine the results of Section 5.5. To start, let us reconsider the definition of $\triangleright^!$.

Definition 6.19 (Refinement of $\triangleright^!$). Let S be a special set of function symbols. \triangleright^S is the relation on base-type terms (and \triangleright^S its reflexive closure) generated by the following clauses:

- $(\lambda x. s) \cdot t_0 \cdots t_n \triangleright^S u$ if $s[x := t_0] \cdot t_1 \cdots t_n \triangleright^S u$
- $f(s_1, \dots, s_m) \cdot t_1 \cdots t_n \triangleright^S u$ if $s_i \cdot \vec{c} \triangleright^S u$ and $f \in S$ \leftarrow here we differ from $\triangleright^!$
- $s \cdot t_1 \cdots t_n \triangleright^S u$ if $t_i \cdot \vec{c} \triangleright^S u$ (s may have any form)

Note that our original definition of $\triangleright^!$ is just a special case of this definition; $\triangleright^!$ can be described as $\triangleright^{\mathcal{F}}$. For local AFSs, we can limit ourselves to $\triangleright^{\mathcal{F}^-}$, shortly denoted \triangleright^- .

In correspondence with Theorem 5.21, we derive the following result:

Theorem 6.20. *A set of dependency pairs $\mathcal{P} = \mathcal{P}_1 \uplus \mathcal{P}_2$ is tagged-chain-free if \mathcal{P}_2 is tagged-chain-free and there is a reduction pair (\succeq, \succ) such that:*

- (1) $\bar{l} \succ \overline{\text{tag}(p)}$ for all $l \rightsquigarrow p \in \mathcal{P}_1$;
- (2) $\bar{l} \succeq \overline{\text{tag}(p)}$ for all $l \rightsquigarrow p \in \mathcal{P}_2$;
- (3) $l \succeq \text{tag}(r)$ for all $l \rightarrow r \in \text{FR}(\mathcal{P})$;
- (4) $f^-(x_1, \dots, x_n) \succeq f(x_1, \dots, x_n)$ for all $f^- \in \mathcal{F}^-$;
- (5) if \mathcal{P} is collapsing, then (\succeq, \succ) respects \triangleright^- , and $f^-(\vec{x}) \succeq f^\sharp(\vec{x})$ for all $f \in \mathcal{D}$.

Proof of Theorem 6.20. Let (\succeq, \succ) be a reduction pair satisfying the requirements in the Theorem, and let $(\succeq, \succeq_1, \succ)$ be the reduction triple generated by (\succeq, \succ) as defined in Section 5.5. This triple clearly satisfies the first four requirements of Theorem 6.18. For the last one, let \mathcal{P} be collapsing; we must see that \succeq has the tagged subterm property.

So let x, s, t, u be given such that $s \triangleright u \triangleright x$. We must see that $(\lambda x. \text{tag}_{\{x\}}(s)) \cdot \text{tag}(t) \cdot \vec{y} (\succ \cup \succeq)^* \text{tag}(u^\sharp[x := t]\gamma) \cdot \vec{v}$ for some substitution γ with $x \notin \text{dom}(\gamma)$ and some terms \vec{v} . Since \succeq contains **beta** and $\text{tag}_{\{x\}}(s \cdot \vec{y}) = \text{tag}_{\{x\}}(s) \cdot \vec{y}$ it suffices if we can prove that for all base-type terms q with $q \triangleright u \triangleright x$ we have: $\text{tag}_{\{x\}}(q)[x := \text{tag}(t)] (\succ \cup \succeq)^* \text{tag}(u^\sharp[x := t]\gamma) \cdot \vec{v}$ for some γ, \vec{v} . This gives what we need by choosing $q = s \cdot \vec{y}$. We prove this statement by induction on q , ordered with \triangleright^- (it is easy to see that this relation is well-founded).

Note (**): $\text{tag}_{\{x\}}(w) \succeq \text{tag}(w)$ by Lemma 6.6 and because always $g^-(\vec{z}) \succeq g(\vec{z})$.

For the base case, let $q = u \cdot v_1 \cdots v_n$ for some terms \vec{v} . Then $\text{tag}_{\{x\}}(q)[x := \text{tag}(t)] = \text{tag}_{\{x\}}(u)[x := \text{tag}(t)] \cdot v'_1 \cdots v'_n$, where each $v'_i = \text{tag}_{\{x\}}(v_i)[x := \text{tag}(t)]$. Since $u \triangleright x$ we know that $\text{tag}_{\{x\}}(u) \succeq \text{tag}(u^\sharp)$: either u does not have the form $f(u_1, \dots, u_m)$ with $f \in \mathcal{D}$, in which case $\text{tag}_{\{x\}}(u) \succeq \text{tag}(u) = \text{tag}(u^\sharp)$ by (**), or u does have this form and $\text{tag}_{\{x\}}(u) = f^-(\text{tag}_{\{x\}}(u_1), \dots, \text{tag}_{\{x\}}(u_m)) \succeq f^-(\text{tag}(u_1), \dots, \text{tag}(u_m)) \succeq f^\sharp(\text{tag}(u_1), \dots, \text{tag}(u_m)) = \text{tag}(u^\sharp)$ by assumption (\mathcal{P} is collapsing, so both properties from (5) hold). Thus, $\text{tag}_{\{x\}}(q)[x := \text{tag}(t)] \succeq \text{tag}(u^\sharp)[x := \text{tag}(t)] \cdot \vec{v}' = \text{tag}(u^\sharp[x := t]) \cdot \vec{v}'$.

Now to consider each of the induction cases:

- (1) $q = (\lambda y. s) \cdot w \cdot v_1 \cdots v_n$ and $s \triangleright u$;

- (2) $q = f(w_1, \dots, w_m) \cdot v_1 \cdots v_n$ and some $w_i \triangleright u$.
 (3) $q = q' \cdot v_1 \cdots v_n$ and one of the $v_i \triangleright u$;

These are the only forms q can have. In very general terms, each of these cases is easy because $\succ \cup \succeq$ includes \triangleright^- (in case 2 we use that $FV(f(\vec{w}))$ contains x , so the tagging function replaces f by f^-). Precise derivations are given below.

In case 1, $\text{tag}_{\{x\}}(q)[x := \text{tag}(t)] \succeq (\text{tag}_{\{x,y\}}(s)[y := \text{tag}_{\{x\}}(w)] \cdot \text{tag}_{\{x\}}(\vec{v}))[x := \text{tag}(t)]$ (since \succeq contains **beta**), $\succeq (\text{tag}_{\{x\}}(s[y := w]) \cdot \text{tag}_{\{x\}}(\vec{v}))[x := \text{tag}(t)]$ by Lemma 6.8, and this is exactly $\text{tag}_{\{x\}}(s[y := w] \cdot \vec{v})[x := \text{tag}(t)]$. Since $s[y := w] \triangleright u[y := w]$ we can use the induction hypothesis to find γ', \vec{v}' such that this term $(\succ \cup \succeq)^* \text{tag}(u[y := w][x := t]\gamma^\#) \cdot \vec{v}' = \text{tag}(u[x := t][y := w[x := t]]\gamma'^\#) \cdot \vec{v}'$, which proves the statement for $\gamma := [y := w[x := t]]\gamma'$.

In case 2, $\text{tag}_{\{x\}}(q) = f^-(\text{tag}_{\{x\}}(\vec{w})) \cdot \text{tag}_{\{x\}}(\vec{v})$ because $f(\vec{w}) \triangleright x$, and this $(\succ \cup \succeq)^* \text{tag}_{\{x\}}(w_i) \cdot \vec{c} = \text{tag}_{\{x\}}(w_i \cdot \vec{c})$ because (\succ, \succeq) respects \triangleright^- . We complete by induction.

Finally, case 3. $\text{tag}_{\{x\}}(q) = \text{tag}_{\{x\}}(q') \cdot \text{tag}_{\{x\}}(v_1) \cdots \text{tag}_{\{x\}}(v_n)$ $(\succ \cup \succeq)^* \text{tag}_{\{x\}}(v_i) \cdot \vec{c} = \text{tag}_{\{x\}}(v_i \cdot \vec{c})$, which by the induction hypothesis $(\succ \cup \succeq)^* \text{tag}(u^\# \gamma[x := t]) \cdot \vec{v}$ as required. \square

Example 6.21. The dependency graph of `twice` has only one SCC, as given in Example 5.13, and whose formative rules we calculated in Example 6.11. Therefore, by Theorems 6.18 and 6.20, `twice` is terminating if there is a reduction pair (\succeq, \succ) which respects \triangleright^- , and orients the following constraints:

$$\begin{array}{llll}
 l^\#(s(n)) & \succ & \text{twice}(\lambda x. l^-(x)) \cdot n & \text{twice}^\#(F) \cdot x \succ F \cdot (F \cdot c_{\text{nat}}) \\
 l^\#(s(n)) & \succ & \text{twice}^\#(\lambda x. l^-(x)) \cdot c_{\text{nat}} & \text{twice}^\#(F) \cdot x \succ F \cdot c_{\text{nat}} \\
 \text{twice}(F) \cdot m & \succ & F \cdot (F \cdot m) & \text{twice}(F) \cdot m \succ F \cdot m \\
 l(s(n)) & \succeq & s(\text{twice}(\lambda x. l^-(x)) \cdot n) & \text{twice}(F) \cdot m \succeq F \cdot (F \cdot m) \\
 f^-(\vec{x}) & \succeq & f(\vec{x}) \ (\forall f \in \mathcal{F}) & f^-(n) \succeq f^\#(n) \ (\forall f \in \mathcal{D})
 \end{array}$$

The top six are requirements for orienting dependency pairs, the next two are the formative rules of this SCC, and the final two are required to deal with the marked symbols.

Theorem 6.20 is a real improvement over Theorem 5.21 because (A) we only need to orient the formative rules of a set of dependency pairs, (B) the requirement that \triangleright^- is included in $\succ \cup \succeq$ is significantly weaker than the requirement for $\triangleright^!$ to be included, and (C) the requirement that $f(\vec{x}) \succeq f^\#(\vec{x})$ was replaced by the requirement that $f^-(\vec{x}) \succeq f^\#(\vec{x})$, which removes the direct relationship between a defined symbol and its marked version. In the next sections we will see how we can use this increased strength.

7. FINDING A REDUCTION PAIR

In this section, we will see two different ways of finding a reduction pair to solve the constraints generated by either the basic or local dependency pair approach. First, we consider how interpretations in a so-called weakly monotonic algebra can be used in the dependency pair setting. Second, we show how to use argument functions to alter an existing reduction pair such as the higher-order recursive path ordering.

Whether we use Theorem 5.21 or Theorem 6.20, we have:

- a set Σ of function symbols: in the basic case, $\Sigma = \mathcal{F}_c^\#$, in the local case $\Sigma = \mathcal{F}_c^\# \cup \{f^- : \sigma \mid f : \sigma \in \mathcal{F}\}$
- a set S of “protected” function symbols: in the basic case, $S = \mathcal{F}$, in the local case $S = \{f^- : \sigma \mid f : \sigma \in \mathcal{F}\}$

- a set A of constraints of the form $l \succ^? p$ and a set B of constraints $l \succeq r$: in the basic case these are given by the dependency pairs and the rules, in the local case the right-hand side is adapted with **tag** and B contains only the formative rules.

In both cases, we must find a reduction pair (\succeq, \succ) such that $l \succ p$ for at least one of the constraints in A , $l \succeq p$ for the remainder of them, and $l \succeq r$ for the constraints in B . Moreover, the reduction pair may have to respect \triangleright^S , which is definitely the case if:

- $s \cdot t_1 \cdots t_n \succeq t_i \cdot \vec{c}$ if both sides have base type;
- $f(s_1, \dots, s_m) \cdot \vec{t} \succeq s_i \cdot \vec{c}$ if $f \in S$ and both sides have base type

We will consider two ways of finding a reduction pair which uses the possibilities created by the dependency pair approach. First, we shall consider *weakly monotonic algebras*, where we natively have a quasi-ordering \succeq which is not just the reflexive closure of \succ , and which is not a simplification ordering. Second, we show how an existing reduction pair can be modified with *argument functions*, a generalisation of the notion of argument filterings.

7.1. Weakly Monotonic Algebras. A semantic method to prove termination of first-order term rewriting is to interpret terms in a well-founded algebra, such that whenever $s \rightarrow t$, their interpretations in the algebra decrease: $\llbracket s \rrbracket > \llbracket t \rrbracket$. Such an algebra is called a *termination model* if $\llbracket l \rrbracket > \llbracket r \rrbracket$ for every rewrite rule $l \rightarrow r$, and some additional properties guarantee that this implies $\llbracket C[l\gamma] \rrbracket > \llbracket C[r\gamma] \rrbracket$ for all contexts C and substitutions γ . A first-order TRS is terminating if and only if it has a termination model [17, 39]. In his PhD thesis [32], van de Pol (extending on a joint paper with Schwichtenberg [33]), generalises this approach to HRSs, with higher-order rewriting modulo $\alpha\beta\eta$, and shows that a HRS is terminating if it has a termination model; the converse does not hold.

Here we consider interpretations of AFS terms in a weakly monotonic algebra, and use these to solve dependency pair constraints. Since $>$ does not have to be monotonic when using dependency pairs, the theory of [32] can be significantly simplified.

Type Interpretation In first-order algebra interpretations, all terms are mapped to an element of some well-founded set \mathcal{A} . In the higher-order setting it turns out to be impractical to map all terms to the same set. Rather, terms of a type $\sigma \Rightarrow \tau$ are mapped to functions.

As a basis, let $\mathcal{A} = (A, \vee, 0, >)^2$, where A is a set, \vee a binary operator on A , 0 an element of A and $>$ a well-founded partial order on A (with reflexive closure \succeq), such that:

- 0 is a minimum element, so $a \geq 0$ for all $a \in A$;
- $x \vee y \geq x$ and $x \vee y \geq y$ for all $x, y \in A$
- $x \vee 0 = x$ for all $x, y \in A$

To each type σ we associate a set WM_σ of *weakly monotonic functionals* and two relations \sqsubset_{wm}^σ and \sqsupset_{wm}^σ , defined inductively as follows.

For a base type ι :

- $\text{WM}_\iota = A$,
- $\sqsubset_{wm}^\iota = >$, and $\sqsupset_{wm}^\iota = \geq$

For a functional type $\sigma \Rightarrow \tau$:

- $\text{WM}_{\sigma \Rightarrow \tau}$ consists of the functions from WM_σ to WM_τ , such that \sqsupset_{wm} is preserved (that is, if $x \sqsupset_{wm}^\sigma y$ then $f(x) \sqsupset_{wm}^\tau f(y)$),

² \mathcal{A} might be a well-ordered monoid or a join-semilattice, but both of these have requirements on \vee which we will not need; with our definitions, \vee may be a supremum operator, but also for instance addition.

- $f \sqsubseteq_{wm}^{\sigma \Rightarrow \tau} g$ iff $f(x) \sqsubseteq_{wm} g(x)$ for all $x \in \text{WM}_\sigma$,
- $f \sqsubseteq_{wm}^{\sigma \Rightarrow \tau} g$ iff $f(x) \sqsubseteq_{wm} g(x)$ for all $x \in \text{WM}_\sigma$.

Thus, $\text{WM}_{\sigma \Rightarrow \tau}$ is a subset of the function space $\text{WM}_\sigma \rightarrow \text{WM}_\tau$, consisting of functions which preserve \sqsubseteq_{wm} . \sqsubseteq_{wm}^σ and \sqsubseteq_{wm}^τ are an order and quasi-order respectively for all types σ , and they are compatible. We usually omit the type σ in the notation, and write just \sqsubseteq_{wm} and \sqsubseteq_{wm} . If either $x \sqsubseteq_{wm} y$ or $x = y$ then $x \sqsubseteq_{wm} y$, but the converse implication does not hold.

This definition differs from the one in [32] in that we use only one set A rather than different sets A_ι for all base types ι ; this is done because we must occasionally compare terms of different types. The original definition also does not use an operator \vee (but does use a \oplus , which satisfies the requirements); this we added because we work with AFSs rather than HRS, and it is convenient to have a “maximum” function to interpret application.

Term Interpretation For some valuable background, let us first consider the most relevant definitions and results from [32]. In the HRS formalism considered there, function symbols do not have an arity; they come equipped with a type, rather than a type declaration. An *interpretation function* \mathcal{J} on the signature is used to associate to each closed term a weakly monotonic functional. Let s be a λ -term, \mathcal{J}_f an element of WM_σ for all $f : \sigma \in \mathcal{F}$, and α a *valuation* which assigns to all variables $x : \sigma \in FV(s)$ an element of WM_σ . Then $[s]_{\mathcal{J}, \alpha}$ is defined by the following clauses:

$$\begin{aligned} [f]_{\mathcal{J}, \alpha} &= \mathcal{J}_f && \text{if } f : \sigma \in \mathcal{F} \\ [x]_{\mathcal{J}, \alpha} &= \alpha(x) && \text{if } x \in \mathcal{V} \\ [s \cdot t]_{\mathcal{J}, \alpha} &= [s]_{\mathcal{J}, \alpha}([t]_{\mathcal{J}, \alpha}) \\ [\lambda x. s]_{\mathcal{J}, \alpha} &= \lambda n. [s]_{\mathcal{J}, \alpha \cup \{x \mapsto n\}} && \text{if } x \notin \text{dom}(\alpha) \text{ (always applicable with } \alpha\text{-conversion)} \end{aligned}$$

Here, λ denotes function construction: $\lambda x_1 \dots x_n. P(x_1, \dots, x_n)$ is the function which takes n arguments x_1, \dots, x_n and returns $P(x_1, \dots, x_n)$.

Lemma 7.1. *Some facts on these interpretations:*

- (1) (Substitution Lemma) *Given a substitution $\gamma = [x_1 := s_1, \dots, x_n := s_n]$ and a valuation α whose domain does not include the x_i : $[s\gamma]_{\mathcal{J}, \alpha} = [s]_{\mathcal{J}, \alpha \circ \gamma}$. Here, $\alpha \circ \gamma$ is the valuation $\alpha \cup \{x_1 \mapsto [s_1]_{\mathcal{J}, \alpha}, \dots, x_n \mapsto [s_n]_{\mathcal{J}, \alpha}\}$.*
- (2) *If $s : \sigma$ is a term, then $[s]_{\mathcal{J}, \alpha} \in \text{WM}_\sigma$ for all valuations α .*

Lemma 7.1(2) provides a convenient way to find weakly monotonic functionals. For example, For $n \in A$ and $\sigma = \tau_1 \Rightarrow \dots \Rightarrow \tau_k \Rightarrow \iota$ the *constant function* $n_\sigma = \lambda x_1 \dots x_k. n \in \text{WM}_\sigma$ (as $n_\sigma = [\lambda x_1 \dots x_k. y]_{\{y \mapsto n\}}$). Similarly, writing $f(\vec{0})$ for $f(0_{\tau_1}, \dots, 0_{\tau_k})$, the function $\lambda f. f(\vec{0})$ is in $\text{WM}_{\sigma \Rightarrow \iota}$. A weakly monotonic functional not given in [32], but which will be needed to deal with term application, is \max_σ , defined as follows:

$$\begin{aligned} \max_\iota(x, y) &= x \vee y && (\text{for } x, y \in A) \\ \max_{\sigma \Rightarrow \tau}(f, y) &= \lambda x. \max_\tau(f(x), y) && (\text{for } f \in \text{WM}_{\sigma \Rightarrow \tau}, y \in \mathcal{A}) \end{aligned}$$

Using induction on the type of the first argument (and once more Lemma 7.1(2)), it is easy to see that $\max_\sigma \in \text{WM}_{\sigma \Rightarrow \iota \Rightarrow \sigma}$.

In HRSs, terms are $\alpha\beta\eta$ -equivalence classes, so $[(\lambda x. s) \cdot t]_{\mathcal{J}, \alpha} = [s]_{\mathcal{J}, \alpha}$ if $x \notin FV(s)$, which is not very convenient in the present setting of AFSs, where terms are considered modulo α only. To adapt the results, we can think of application as a function symbol.

Definition 7.2. A *signature interpretation* associates a weakly monotonic functional $\mathcal{J}_f \in \text{WM}_{\sigma_1 \Rightarrow \dots \Rightarrow \sigma_n \Rightarrow \tau}$ to every function symbol $f : [\sigma_1 \times \dots \times \sigma_n] \Rightarrow \tau$ of the signature. The pair

$(\mathcal{A}, \mathcal{J})$ is a *weakly monotonic algebra*. A *valuation* is a function α with a finite domain of variables, such that $\alpha(x) \in \mathbf{WM}_\sigma$ for every variable $x : \sigma$ in its domain.

Fixing \mathcal{J} and α , the *interpretation* of a term s , denoted $\llbracket s \rrbracket_{\mathcal{J}, \alpha}$, is a weakly monotonic functional defined by induction on the definition of terms as follows:

$$\begin{aligned} \llbracket x \rrbracket_{\mathcal{J}, \alpha} &= \alpha(x) \text{ if } x \in \mathcal{V} \\ \llbracket f(s_1, \dots, s_n) \rrbracket_{\mathcal{J}, \alpha} &= \mathcal{J}_f(\llbracket s_1 \rrbracket_{\mathcal{J}, \alpha}, \dots, \llbracket s_n \rrbracket_{\mathcal{J}, \alpha}) \\ \llbracket \lambda x. s \rrbracket_{\mathcal{J}, \alpha} &= \lambda n. \llbracket s \rrbracket_{\mathcal{J}, \alpha \cup \{x \mapsto n\}} \text{ if } x \notin \text{dom}(\alpha) \\ \llbracket s \cdot t \rrbracket_{\mathcal{J}, \alpha} &= \max(\llbracket s \rrbracket_{\mathcal{J}, \alpha}(\llbracket t \rrbracket_{\mathcal{J}, \alpha}), \llbracket t \rrbracket_{\mathcal{J}, \alpha}(\vec{0})) \end{aligned}$$

We assume α is defined on all free variables of s . This definition corresponds with the one for HRSs, if we replace application by a function symbol which is interpreted with \max .

Example 7.3. In our running example twice, consider an interpretation in the natural numbers. Say $\mathcal{J}_1 = \lambda n. n$ and $\mathcal{J}_s = \lambda n. n + 1$. Then $\llbracket l(s(x)) \rrbracket_{\mathcal{J}, \alpha} = \alpha(x) + 1$.

Reduction Pair. Since \sqsubset_{wm} is in general not closed under taking contexts, it cannot be used directly like in first-order rewriting: $\llbracket l \rrbracket_{\mathcal{J}, \alpha} \sqsubset_{wm} \llbracket r \rrbracket_{\mathcal{J}, \alpha}$ does not in general imply $\llbracket C[l\gamma] \rrbracket_{\mathcal{J}, \alpha} \sqsubset_{wm} \llbracket C[r\gamma] \rrbracket_{\mathcal{J}, \alpha}$. This issue (which Van De Pol handles by imposing the restriction that \mathcal{J}_f must be *strict*) disappears entirely in the context of dependency pairs.

Theorem 7.4. Let \mathcal{J} be an interpretation of the signature Σ^3 such that:

- \mathcal{J} maps each c_σ to 0_σ ;
- for all $f : [\sigma_1 \times \dots \times \sigma_m] \Rightarrow \tau_1 \Rightarrow \dots \Rightarrow \tau_k \Rightarrow \iota \in S$, all $1 \leq i \leq m$ and all $n \in \mathbf{WM}_{\sigma_i}$:
 $\mathcal{J}_f(0_{\sigma_1}, \dots, n, \dots, 0_{\sigma_m}, 0_{\tau_1}, \dots, 0_{\tau_k}) \sqsupseteq_{wm} n(\vec{0})$.

Define $s \succeq t$ if $\llbracket s \rrbracket_{\mathcal{J}, \alpha} \sqsupseteq_{wm} \llbracket t \rrbracket_{\mathcal{J}, \alpha}$ for all valuations α , and $s \succ t$ if $\llbracket s \rrbracket_{\mathcal{J}, \alpha} \sqsubset_{wm} \llbracket t \rrbracket_{\mathcal{J}, \alpha}$ for all valuations α . Then (\succeq, \succ) is a reduction pair which respects \triangleright^S .

Proof. It is easy to see that \succ is an ordering and \succeq a quasi-ordering; \succ is well-founded because if $f \sqsubset_{wm} g$ then also $f(\vec{0}) \sqsubset_{wm} g(\vec{0})$ (and \sqsubset_{wm} is well-founded in \mathcal{A}). \succeq is monotonic as we see with a simple case distinction; if $\llbracket s \rrbracket_{\mathcal{J}, \alpha} \sqsupseteq_{wm} \llbracket t \rrbracket_{\mathcal{J}, \alpha}$ for all α , then:

- $\llbracket \lambda x. s \rrbracket_{\mathcal{J}, \alpha} = \lambda n. \llbracket s \rrbracket_{\mathcal{J}, \alpha \cup \{x \mapsto n\}} \sqsupseteq_{wm} \lambda n. \llbracket t \rrbracket_{\mathcal{J}, \alpha \cup \{x \mapsto n\}} = \llbracket \lambda x. t \rrbracket_{\mathcal{J}, \alpha}$
- $\llbracket s \cdot u \rrbracket_{\mathcal{J}, \alpha} \sqsupseteq_{wm} \llbracket t \cdot u \rrbracket_{\mathcal{J}, \alpha}$ by weak monotonicity of $\lambda f. \lambda n. \max(f(n), n(\vec{0}))$
- $\llbracket u \cdot s \rrbracket_{\mathcal{J}, \alpha} \sqsupseteq_{wm} \llbracket u \cdot t \rrbracket_{\mathcal{J}, \alpha}$ by weak monotonicity of $\lambda f. \lambda n. \max(f(n), n(\vec{0}))$
- $\llbracket f(\dots, s, \dots) \rrbracket_{\mathcal{J}, \alpha} \sqsupseteq_{wm} \llbracket f(\dots, t, \dots) \rrbracket_{\mathcal{J}, \alpha}$ by weak monotonicity of \mathcal{J}_f

In addition, \succeq contains **beta** because $\llbracket (\lambda x. s) \cdot t \rrbracket_{\mathcal{J}, \alpha} = \max(\llbracket s \rrbracket_{\mathcal{J}, \alpha \circ [x \mapsto \llbracket t \rrbracket_{\mathcal{J}, \alpha}]}(\llbracket t \rrbracket_{\mathcal{J}, \alpha}(\vec{0})), \llbracket t \rrbracket_{\mathcal{J}, \alpha}(\vec{0})) \sqsupseteq_{wm} \llbracket s \rrbracket_{\mathcal{J}, \alpha \circ [x \mapsto \llbracket t \rrbracket_{\mathcal{J}, \alpha}]}(\llbracket t \rrbracket_{\mathcal{J}, \alpha}(\vec{0}))$, which is exactly $\llbracket s[x := t] \rrbracket_{\mathcal{J}, \alpha}$ by the substitution Lemma. Compatibility is inherited from compatibility of \sqsubset_{wm} and \sqsupseteq_{wm} on \mathcal{A} (where \sqsupseteq_{wm} is the reflexive closure of \sqsubset_{wm}): if $s \succ t \succeq u$, then $\llbracket s \rrbracket_{\mathcal{J}, \alpha} = \lambda n_1 \dots n_k. f(\vec{n}) \sqsubset_{wm} \llbracket t \rrbracket_{\mathcal{J}, \alpha} = \lambda \vec{n}. g(\vec{n}) \sqsupseteq_{wm} \llbracket u \rrbracket_{\mathcal{J}, \alpha} = \lambda \vec{n}. h(\vec{n})$, and we are done because also $f(\vec{n}) \sqsubset_{wm} h(\vec{n})$.

Finally, stability follows by the substitution Lemma: always $\llbracket u\gamma \rrbracket_{\mathcal{J}, \alpha} = \llbracket u \rrbracket_{\mathcal{J}, \alpha \circ \gamma}$, so if $\llbracket s \rrbracket_{\mathcal{J}, \alpha} \sqsubset_{wm} \llbracket t \rrbracket_{\mathcal{J}, \alpha}$ for all valuations α , this also holds for the valuation $\alpha' = \alpha \circ \gamma$.

As observed at the start of Section 7, (\succeq, \succ) respects \triangleright^S if:

- $\llbracket s \cdot \vec{t} \rrbracket_{\mathcal{J}, \alpha} \sqsupseteq_{wm} \llbracket t_i \cdot \vec{c} \rrbracket_{\mathcal{J}, \alpha}$; this holds because, by the use of \max for applications, $\llbracket s \cdot \vec{t} \rrbracket_{\mathcal{J}, \alpha} \sqsupseteq_{wm} \llbracket t_i \rrbracket_{\mathcal{J}, \alpha}(\vec{0}) = \max(\dots \max(\llbracket t_i \rrbracket_{\mathcal{J}, \alpha}(\vec{0}), 0_{\sigma_1}), \dots, 0_{\sigma_n}) = \llbracket t_i \cdot \vec{c} \rrbracket_{\mathcal{J}, \alpha}$;

³ Σ is the signature introduced at the start of Section 7, either $\mathcal{F}_c^\#$ or $\mathcal{F}_c^\# \cup \{f^- : \sigma \mid f : \sigma \in \mathcal{F}\}$.

- $\llbracket f(\vec{s}) \cdot \vec{t} \rrbracket_{\mathcal{J}, \alpha} \sqsupseteq_{wm} \llbracket s_i \cdot \vec{c} \rrbracket_{\mathcal{J}, \alpha}$ if $f \in S$; this holds because $\llbracket f(\vec{s}) \cdot \vec{t} \rrbracket_{\mathcal{J}, \alpha} \sqsupseteq_{wm} \mathcal{J}_f(\llbracket s_1 \rrbracket_{\mathcal{J}, \alpha}, \dots, \llbracket s_n \rrbracket_{\mathcal{J}, \alpha}, \llbracket t_1 \rrbracket_{\mathcal{J}, \alpha}, \dots, \llbracket t_m \rrbracket_{\mathcal{J}, \alpha})$, which by weak monotonicity and assumption $\sqsupseteq_{wm} \mathcal{J}_f(0_{\sigma_1}, \dots, \llbracket s_i \rrbracket_{\mathcal{J}, \alpha}, \dots, 0_{\sigma_n}, 0_{\tau_1}, \dots, 0_{\tau_m}) \sqsupseteq_{wm} \llbracket s_i \rrbracket_{\mathcal{J}, \alpha}(\vec{0}) = \llbracket s_i \cdot \vec{c} \rrbracket_{\mathcal{J}, \alpha}$. \square

Although it is not in general possible to (automatically) determine whether a suitable interpretation exists, one could for instance try polynomial interpretations over the natural numbers. Since addition and multiplication are both weakly monotonic, such an interpretation is sound, and like in first-order rewriting there are some easily automatable techniques for finding suitable polynomials. The automation of individual reduction pairs is beyond the scope of this paper, however; we refer to [9] for a more detailed discussion.

Example 7.5. To prove that $\mathcal{C}_{\text{twice}}$ is tagged-chain-free, write $\mathcal{C}_{\text{twice}} = \mathcal{P}_1 \uplus \mathcal{P}_2$, where \mathcal{P}_1 consists of the two $I^\#$ dependency pairs, and \mathcal{P}_2 of the remainder. By Theorem 6.18 it suffices to prove that \mathcal{P}_2 is tagged-chain-free, if we can find a reduction pair which respects \triangleright^- and orients the following requirements:

$$\begin{array}{llll}
I^\#(s(n)) & \succ & \text{twice}(\lambda x. I^-(x)) \cdot n & \text{twice}^\#(F) \cdot x \succeq F \cdot (F \cdot c_{\text{nat}}) \\
I^\#(s(n)) & \succ & \text{twice}^\#(\lambda x. I^-(x)) \cdot c_{\text{nat}} & \text{twice}^\#(F) \cdot x \succeq F \cdot c_{\text{nat}} \\
\text{twice}(F) \cdot m & \succeq & F \cdot (F \cdot m) & \text{twice}(F) \cdot m \succeq F \cdot m \\
I(s(n)) & \succeq & s(\text{twice}(\lambda x. I^-(x)) \cdot n) & \text{twice}(F) \cdot m \succeq F \cdot (F \cdot m) \\
f^-(\vec{x}) & \succeq & f(\vec{x}) \ (\forall f \in \mathcal{F}) & f^-(n) \succeq f^\#(n) \ (\forall f \in \mathcal{D})
\end{array}$$

Let:

- $\mathcal{J}_I = \mathcal{J}_{I^\#} = \mathcal{J}_{I^-} = \lambda n. n$
- $\mathcal{J}_O = \mathcal{J}_{O^-} = 0$
- $\mathcal{J}_s = \mathcal{J}_{s^-} = \lambda n. n$
- $\mathcal{J}_{\text{twice}} = \mathcal{J}_{\text{twice}^\#} = \mathcal{J}_{\text{twice}^-} = \lambda f n. f(f(n))$

Then it is clear that the bottom two constraints are satisfied. With some calculation we can see that the others hold as well, using the following case distinction for the `twice` cases: if $m \geq F(m)$, then also $m \geq F(F(m))$ by weak monotonicity; if $F(m) \geq m$, then also $F(F(m)) \geq m$. Therefore $\max(F(F(m)), m) \geq \max(F(\max(F(m), m)), \max(F(m), m))$.

To complete the termination proof of `twice`, it suffices to find a reduction pair such that:

$$\begin{array}{ll}
\text{twice}(F) \cdot n \succ F \cdot (F \cdot n) & \text{twice}^\#(F) \cdot n \succ F \cdot (F \cdot c_{\text{nat}}) \\
\text{twice}(F) \cdot n \succ F \cdot n & \text{twice}^\#(F) \cdot n \succ F \cdot c_{\text{nat}}
\end{array}$$

This is satisfied with an interpretation with $\mathcal{J}_{\text{twice}^\#} = \mathcal{J}_{\text{twice}} = \lambda f n. \max(f(f(n)), n) + 1$.

Example 7.6. A well-known example of higher-order rewriting is `map`:

$$\begin{array}{ll}
\text{map}(F, \text{nil}) & \rightarrow \text{nil} \\
\text{map}(F, \text{cons}(h, t)) & \rightarrow \text{cons}(F \cdot h, \text{map}(F, t))
\end{array}$$

Applying Theorem 6.18 to prove termination, it suffices to find a reduction pair (\succeq, \succ) with:

$$\begin{array}{ll}
\text{map}^\#(F, \text{cons}(h, t)) & \succ F \cdot h \\
\text{map}^\#(F, \text{cons}(h, t)) & \succ \text{map}^\#(F, t) \\
\text{map}(F, \text{cons}(h, t)) & \succeq \text{cons}(F \cdot h, \text{map}(F, t))
\end{array}$$

Note that the elements of S do not occur in the rules, so we can pretty much ignore them (see also Section 9). Using an interpretation in the natural numbers with the usual greater than, 0 and max-operator, consider the following constant interpretation: $\mathcal{J}_{\text{map}^\#} = \lambda f. \lambda x. f(x) +$

x , $\mathcal{J}_{\text{map}} = \lambda f. \lambda x. x \cdot f(x) + x$, $\mathcal{J}_{\text{cons}} = \lambda x. \lambda y. x + y + 1$. Taking $\alpha = [F := f, h := n, t := m]$, the requirements above become:

$$\begin{aligned} f(n + m + 1) + n + m + 1 &> \max(f(n), n) \\ f(n + m + 1) + n + m + 1 &> f(m) + m \\ (n + m + 1) \cdot f(n + m + 1) + n + m + 1 &\geq \max(f(n), n) + m \cdot f(m) + m + 1 \end{aligned}$$

Taking into account that f must be a weakly monotonic functional, we of course have $f(n + m + 1) \geq f(n), f(m)$. Thus, it is not hard to see that all requirements are true.

7.2. Argument Functions. In Section 4.4 we saw that, in the first-order dependency pair approach, (simplification) orderings may be combined with *argument filterings*. Such filterings either eliminate some direct arguments s_i from a term $f(s_1, \dots, s_n)$, or replace the term by one of the s_i . We consider an extension of this technique, called *argument functions*.

Let Σ and S be sets of function symbols, as introduced at the start of Section 7.

Definition 7.7 (Argument Function). Let Σ^{ex} be a set of function symbols, and π a type-respecting function mapping terms $f(x_1, \dots, x_n)$ with $f \in \Sigma$ to some term over Σ^{ex} ; we require that $FV(\pi(f(\vec{x}))) \subseteq \{\vec{x}\}$. We extend π to a function $\bar{\pi}(\cdot)$, called an *argument function*, on all terms as follows:

$$\begin{aligned} \bar{\pi}(s \cdot t) &= \bar{\pi}(s) \cdot \bar{\pi}(t) \\ \bar{\pi}(\lambda x. s) &= \lambda x. \bar{\pi}(s) \\ \bar{\pi}(x) &= x \quad (x \in \mathcal{V}) \\ \bar{\pi}(f(s_1, \dots, s_n)) &= \pi(f(x_1, \dots, x_n))[x_1 := \bar{\pi}(s_1), \dots, x_n := \bar{\pi}(s_n)] \end{aligned}$$

An *argument filtering* is an argument function where each $\pi(f(x_1, \dots, x_n))$ has the form $f'(x_{i_1}, \dots, x_{i_k})$, or x_i . However, we can choose π entirely different as well. For example, if $\pi(\text{twice}(x)) = \lambda y. x \cdot y$, then the term $\text{twice}(\lambda n. l(n)) \cdot o$ is mapped to $(\lambda y. (\lambda n. l(n)) \cdot y) \cdot o$.

An argument function π has the *S-subterm property* if $FV(\pi(f(\vec{x}))) = \{\vec{x}\}$ for all $f \in S$.

Theorem 7.8. Let (\succeq, \succ) be a reduction pair on terms over Σ^{ex} . Define $\geq, >$ on terms over Σ as follows: $s \geq t$ iff $\bar{\pi}(s) \succeq \bar{\pi}(t)$ and $s > t$ iff $\bar{\pi}(s) \succ \bar{\pi}(t)$.

Then $(\geq, >)$ is a reduction pair. If π has the S-subterm property and (\succeq, \succ) respects $\triangleright^!$ and $s \succeq c_\sigma$ for all terms s of type σ , then $(\geq, >)$ respects \triangleright^S .

Proof. We first make the following observation (**): $\bar{\pi}(s\gamma) = \bar{\pi}(s)\gamma^\pi$, where $\gamma^\pi(x) = \bar{\pi}(\gamma(x))$. This holds by induction on the form of s . The only non-obvious case is when $s = f(s_1, \dots, s_n)$ and $\pi(f(x_1, \dots, x_n)) = p$; then $\bar{\pi}(s\gamma) = p[x_1 := \bar{\pi}(s_1\gamma), \dots, x_n := \bar{\pi}(s_n\gamma)] = p[x_1 := \bar{\pi}(s_1)\gamma^\pi, \dots, x_n := \bar{\pi}(s_n)\gamma^\pi]$ by the induction hypothesis, $= \bar{\pi}(s)\gamma^\pi$ because $FV(p) \subseteq \{\vec{x}\}$.

Having this, it is easy to see that \geq and $>$ are both stable, and compatibility, well-foundedness, transitivity and (anti-)reflexivity are inherited from the corresponding properties of \succeq and \succ . As for monotonicity of \geq , the only non-trivial question is whether $f(s_1, \dots, s_n) \geq f(s'_1, \dots, s'_n)$ when each $s_i \geq s'_i$, but this is clear by monotonicity of \succeq : if each $\bar{\pi}(s_i) \succeq \bar{\pi}(s'_i)$, then $\bar{\pi}(f(\vec{s})) = \pi(f(\vec{x}))[\vec{x} := \bar{\pi}(\vec{s})] \succeq \pi(f(\vec{x}))[\vec{x} := \bar{\pi}(\vec{s}')] = \bar{\pi}(f(\vec{s}'))$. Also \geq contains **beta** by (**): $\bar{\pi}((\lambda x. s) \cdot t) = (\lambda x. \bar{\pi}(s)) \cdot \bar{\pi}(t) \succeq \bar{\pi}(s)[x := \bar{\pi}(t)]$ (as \succeq contains **beta**), and this equals $\bar{\pi}(s[x := t])$ as required. Thus, $(\geq, >)$ is a reduction pair.

As observed at the start of Section 7, to see that $(\geq, >)$ respects \triangleright^S it suffices if $s \cdot t_1 \cdots t_n \geq t_i \cdot \vec{c}$ and $f(s_1, \dots, s_m) \cdot t_1 \cdots t_n \geq s_i \cdot \vec{c}$ if both sides have base type, and $f \in S$. Suppose (\succeq, \succ) respects $\triangleright^!$ and π has the S-subterm property. Then obviously $\bar{\pi}(s \cdot t_1 \cdots t_n) = \bar{\pi}(s) \cdot \bar{\pi}(t_1) \cdots \bar{\pi}(t_n) \succeq \bar{\pi}(t_i) \cdot \vec{c} = \bar{\pi}(t_i \cdot \vec{c})$.

For the second part, let $f \in S$ and $q = f(\vec{s}) \cdot \vec{t}$; we must see that $\bar{\pi}(q) = \pi(f(x_1, \dots, x_m)[x_1 := \bar{\pi}(s_1), \dots, x_m := \bar{\pi}(s_m)] \cdot \bar{\pi}(t_1) \cdots \bar{\pi}(t_n)) \succeq \bar{\pi}(s_i \cdot \vec{c})$. Since π has the S-subterm property, x_i occurs in $\pi(f(\vec{x}))$. Therefore, $\bar{\pi}(s_i)$ is a subterm of $\bar{\pi}(q)$ and it contains no free variables which are bound in $\bar{\pi}(q)$. But then we must have $\bar{\pi}(f(\vec{s}) \cdot \vec{t}) \triangleright^! \bar{\pi}(s_i) \cdot \vec{u}$ for some terms \vec{u} ; since the c_σ are minimal elements by assumption, this term $\succeq \bar{\pi}(s_i) \cdot \vec{c} = \bar{\pi}(u_i \cdot \vec{c})$ as required. \square

Theorem 7.8 allows us to modify terms before applying an ordering. Even argument functions which respect the \mathcal{F} -subterm property can be useful; for example, if there is a rule $f(x_1, \dots, x_n) \rightarrow r$, then an argument function with $\pi(f(\vec{x})) = r$ is probably a good idea.

Using Argument Functions With CPO Unfortunately, Theorem 7.8 cannot directly be used with the higher-order recursive path ordering or the more powerful computability path ordering: these relations are not transitive, do not respect $\triangleright^!$, and do not have minimal elements. But we *can* alter either relation a bit, and obtain a usable reduction pair.

Let $>_{\text{CPO}}$ be the relation given by CPO [4], with some fixed status function, type ordering, and precedence \geq_F on the symbols in $\Sigma \setminus \mathbb{C}$. Introducing new symbols $@^{\sigma, \tau}, \Lambda^{\sigma, \tau}$ and c_σ for all types, extend the precedence with $f >_F @^{\sigma, \tau}, \Lambda^{\sigma, \tau}, c_\sigma$ if $f \in \Sigma \setminus \mathbb{C}$, for all types $@^{\sigma, \tau} >_F \Lambda^{\sigma', \tau'} >_F c_\rho$ and moreover $@^{\sigma, \tau} >_F @^{\rho, \alpha}$ if $\rho \Rightarrow \alpha$ is a strict subtype of $\sigma \Rightarrow \tau$. Define $(\succeq_{\text{CPO}}, \succ_{\text{CPO}})$ as follows: $s \succeq_{\text{CPO}} t$ if for all closed substitutions γ : $\mu(s\gamma) >_{\text{CPO}}^* \mu(t\gamma)$, and $s \succ_{\text{CPO}} t$ if for all closed substitutions γ : $\mu(s\gamma) >_{\text{CPO}}^+ \mu(t\gamma)$. Here, μ is given by:

$$\begin{aligned} \mu(x) &= x && \text{if } x \in \mathcal{V} \\ \mu(\lambda x. s) &= \Lambda^{\sigma, \tau}(\lambda x. \mu(s)) && \text{if } \lambda x. s : \sigma \Rightarrow \tau \\ \mu(s \cdot t) &= @^{\sigma, \tau}(\mu(s), \mu(t)) && \text{if } s : \sigma \Rightarrow \tau \\ \mu(f(s_1, \dots, s_n)) &= f(\mu(s_1), \dots, \mu(s_n)) \end{aligned}$$

Since $\mu(s)[x := \mu(t)] = \mu(s[x := t])$ (as demonstrated by an easy induction), it is not hard to see that $(\succeq_{\text{CPO}}, \succ_{\text{CPO}})$ is indeed a reduction pair, that the c_σ are minimal elements, and that we can ignore the γ : since $>_{\text{CPO}}$ is stable, $s \succ_{\text{CPO}} t$ holds if $\mu(s) >_{\text{CPO}}^+ \mu(t)$.

Moreover, always $s \cdot t_1 \cdots t_n \succeq_{\text{CPO}} t_i \cdot \vec{c}$ and $f(s_1, \dots, s_n) \cdot \vec{t} \succeq_{\text{CPO}} s_i \cdot \vec{c}$:

- In CPO, $g(u, v) >_{\text{CPO}} w$ if $u >_{\text{CPO}} w$, regardless of type differences. Since the $@^{\sigma, \tau}$ are function symbols, we thus have: $\mu(u \cdot v) >_{\text{CPO}} w$ if $\mu(u) >_{\text{CPO}} w$.

Let $t_i : \sigma = \sigma_1 \Rightarrow \dots \Rightarrow \sigma_n \Rightarrow \iota$. Then $\mu(s \cdot \vec{t}) >_{\text{CPO}} \mu(t_i \cdot \vec{c})$ if (selecting the first argument $n - i$ times) $\mu(s \cdot t_1 \cdots t_i) = @^{\sigma, \tau}(@(\dots @(\mu(s), \mu(t_1)), \dots, \mu(t_{i-1})), \mu(t_i)) >_{\text{CPO}} \mu(t_i \cdot \vec{c})$. Since σ is a strict subtype of $\sigma \Rightarrow \tau$, so $@^{\sigma, \tau} >_F @^{\sigma_j, \sigma_{j+1} \Rightarrow \dots \Rightarrow \sigma_n \Rightarrow \iota}$ for all j , and because certainly $@^{\sigma, \tau}(\dots) >_{\text{CPO}} c_{\sigma_j}$ for all j , this indeed holds, as $\mu(t_i \cdot \vec{c}) = @^{\sigma_n, \iota}(\dots @^{\sigma_1, \sigma_2 \Rightarrow \dots \Rightarrow \sigma_n \Rightarrow \iota}(\mu(t_i), c_{\sigma_1}), \dots, c_{\sigma_n})$.

- Similarly, $\mu(f(\vec{s}) \cdot \vec{t}) >_{\text{CPO}} \mu(s_i \cdot \vec{c})$ if $f(\mu(s_1), \dots, \mu(s_n)) >_{\text{CPO}} \mu(s_i \cdot \vec{c})$, which holds because $f >_F @^{\sigma_j, \sigma_{j+1} \Rightarrow \dots \Rightarrow \sigma_n \Rightarrow \iota}$ for all j , and also $f >_F c_{\sigma_j}$.

Rather than altering existing definitions of HORPO and CPO with a transformation like the one above, we might consider definitions which natively have minimal elements – this is an addition which should not be hard to include in the well-foundedness proof.

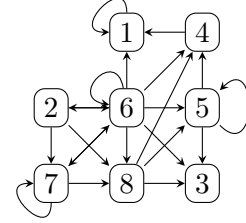
Example 7.9. Consider the following AFS:

$$\mathcal{F} = \left\{ \begin{array}{ll} \text{o} : \text{nat} & \text{if} : [\text{bool} \times \text{list} \times \text{list}] \Rightarrow \text{list} \\ \text{s} : [\text{nat}] \Rightarrow \text{nat} & \text{lteq} : [\text{nat} \times \text{nat}] \Rightarrow \text{bool} \\ \text{true} : \text{bool} & \text{from} : [\text{nat} \times \text{list}] \Rightarrow \text{list} \\ \text{false} : \text{bool} & \text{chain} : [\text{nat} \Rightarrow \text{nat} \times \text{list}] \Rightarrow \text{list} \\ \text{nil} : \text{list} & \text{incch} : [\text{list}] \Rightarrow \text{list} \\ \text{cons} : [\text{nat} \times \text{list}] \Rightarrow \text{list} \end{array} \right\}$$

$$\mathcal{R} = \left\{ \begin{array}{lll} \text{lteq}(\text{s}(x), \text{o}) \rightarrow \text{false} & \text{if}(\text{true}, x, y) \rightarrow x & \\ \text{lteq}(\text{o}, x) \rightarrow \text{true} & \text{if}(\text{false}, x, y) \rightarrow y & \\ \text{lteq}(\text{s}(x), \text{s}(y)) \rightarrow \text{lteq}(x, y) & \text{from}(x, \text{nil}) \rightarrow \text{nil} & \\ \text{incch}(x) \rightarrow \text{chain}(\lambda y. \text{s}(y), x) & \text{chain}(F, \text{nil}) \rightarrow \text{nil} & \\ \text{from}(x, \text{cons}(y, z)) \rightarrow \text{if}(\text{lteq}(x, y), \text{cons}(y, z), \text{from}(x, z)) & & \\ \text{chain}(F, \text{cons}(y, z)) \rightarrow \text{cons}(F \cdot y, \text{chain}(F, \text{from}(F \cdot y, z))) & & \end{array} \right\}$$

This AFS, which appears in the termination problem database (see [38]), has eight dependency pairs, and the following dependency graph:

- (1) $\text{lteq}^\#(\text{s}(x), \text{s}(y)) \rightsquigarrow \text{lteq}^\#(x, y)$
- (2) $\text{incch}^\#(x) \rightsquigarrow \text{chain}^\#(\lambda y. \text{s}(y), x)$
- (3) $\text{from}^\#(x, \text{cons}(y, z)) \rightsquigarrow \text{if}^\#(\text{lteq}(x, y), \text{cons}(y, z), \text{from}(x, z))$
- (4) $\text{from}^\#(x, \text{cons}(y, z)) \rightsquigarrow \text{lteq}^\#(x, y)$
- (5) $\text{from}^\#(x, \text{cons}(y, z)) \rightsquigarrow \text{from}^\#(x, z)$
- (6) $\text{chain}^\#(F, \text{cons}(y, z)) \rightsquigarrow F \cdot y$
- (7) $\text{chain}^\#(F, \text{cons}(y, z)) \rightsquigarrow \text{chain}^\#(F, \text{from}(F \cdot y, z))$
- (8) $\text{chain}^\#(F, \text{cons}(y, z)) \rightsquigarrow \text{from}^\#(F \cdot y, z)$



We consider the SCC consisting of pairs 2, 6 and 7. As this is a local AFS, we may use almost unrestricted argument functions. Let $\pi(\text{s}(x)) = \pi(\text{s}^-(x)) = x$, $\pi(\text{incch}^\#(x)) = \text{chain}^\#(\lambda y. y, x)$, $\pi(\text{incch}(x)) = \text{chain}(\lambda y. y, x)$, $\pi(\text{lteq}(x, y)) = \text{lteq}'$ and $\pi(\text{from})(x, y) = \text{from}'(y)$; for other symbols, π is the identity. Then the three dependency pairs of interest are oriented with $(\succeq_{\text{CPO}}, \succ_{\text{CPO}})$ if $\text{cons} >_F \text{from}'$:

- (2) $\bar{\pi}(\text{incch}^\#(x)) = \text{chain}^\#(\lambda y. y, x) \succeq_{\text{CPO}} \text{chain}^\#(\lambda y. y, x) = \bar{\pi}(\text{chain}^\#(\lambda y. \text{s}^-(y), x))$
- (6) $\text{chain}^\#(F, \text{cons}(y, z)) \succ_{\text{CPO}} F \cdot y$
- (7) $\text{chain}^\#(F, \text{cons}(y, z)) \succ_{\text{CPO}} \text{chain}^\#(F, \text{from}'(z))$

All rules, even the non-formative ones, are oriented as well, with for instance the precedence $\text{chain}, \text{chain}^\# >_F \text{cons} >_F \text{from}' >_F \text{if}, \text{lteq}' >_F \text{false}, \text{true}$

$$\begin{array}{lll} \text{lteq}' & \succeq_{\text{CPO}} & \text{false} & \text{if}(\text{true}, x, y) & \succeq_{\text{CPO}} & x \\ \text{lteq}' & \succeq_{\text{CPO}} & \text{true} & \text{if}(\text{false}, x, y) & \succeq_{\text{CPO}} & y \\ \text{lteq}' & \succeq_{\text{CPO}} & \text{lteq}' & \text{from}'(\text{nil}) & \succeq_{\text{CPO}} & \text{nil} \\ \text{chain}(\lambda y. y, x) & \succeq_{\text{CPO}} & \text{chain}(\lambda y. y, x) & \text{chain}(F, \text{nil}) & \succeq_{\text{CPO}} & \text{nil} \\ \text{from}'(\text{cons}(y, z)) & \succeq_{\text{CPO}} & \text{if}(\text{lteq}', \text{cons}(y, z), \text{from}'(z)) & & & \\ \text{chain}(F, \text{cons}(y, z)) & \succeq_{\text{CPO}} & \text{cons}(F \cdot y, \text{chain}(F, \text{from}'(z))) & & & \end{array}$$

The remaining dependency pair (2) is clearly tagged-chain-free, so this AFS is terminating if the two SCCs $\{(1)\}$ and $\{(5)\}$ are tagged-chain-free.

8. NON-COLLAPSING DEPENDENCY PAIRS

Many powerful aspects of the first-order dependency pair framework, such as the subterm criterion and usable rules, break in the presence of collapsing dependency pairs. But for *parts* of a termination proof, we may have non-collapsing dependency pairs. Consider for example the system from Example 7.9: the dependency graph has three SCCs, and two of them (the SCCs $\{(1)\}$ and $\{(5)\}$) are non-collapsing. Their chain-freeness can be demonstrated with a reduction triple which does not have the limited subterm property.

In this section, we will briefly discuss the subterm criterion and usable rules, limited to non-collapsing sets \mathcal{P} . To avoid double work between the basic and tagged dependency pair approach, we note that tags add very little advantage in this setting: tagging mainly plays a role in the limited or tagged subterm property. The primary advantage of Theorem 6.20 over Theorem 5.21, therefore, is the addition of formative rules. So let us revise the definition of chain-free: *a set \mathcal{P} of dependency pairs is chain-free if there is no minimal dependency chain $[(\rho_i, s_i, t_i) \mid i \in \mathbb{N}]$ with all $\rho_i \in \mathcal{P} \cup \{\mathbf{beta}\}$ such that $t_i \rightarrow_{FR(\mathcal{P})}^* s_{i+1}$ for all i .* Here, $FR(\mathcal{P}) = \mathcal{R}$ if the AFS under consideration is not local.

Lemma 8.1. *A set \mathcal{P} of dependency pairs is tagged-chain-free if it is chain-free.*

Proof. Obvious consequence of Theorem 6.15, by removing the tags. \square

We will consider only how to prove chain-freeness for the sets under consideration.

8.1. The Subterm Criterion. Let \mathcal{P} be non-collapsing, and let \mathcal{H} be the set of function symbols f such that a left- or right-hand side of a dependency pair has the form $f(\vec{s}) \cdot \vec{t}$. A *projection function* for \mathcal{H} is a function ν which assigns to each $f \in \mathcal{H}$ a number i such that for all $l \rightsquigarrow p \in \mathcal{P}$, the following function $\bar{\nu}$ is well-defined for both l and p :

$$\bar{\nu}(f(s_1, \dots, s_m) \cdot s_{m+1} \cdots s_n) = s_{\nu(f)}$$

This differs from the first-order definition only in that we cater for dependency pairs which do not have the form $f^\sharp(l_1, \dots, l_n) \rightsquigarrow g^\sharp(p_1, \dots, p_m)$.

Theorem 8.2. *Let $\mathcal{P} = \mathcal{P}_1 \uplus \mathcal{P}_2$ be a set of non-collapsing dependency pairs, and suppose a projection function ν exists such that $\bar{\nu}(l) \triangleright \bar{\nu}(p)$ for all $l \rightsquigarrow p \in \mathcal{P}_1$ and $\bar{\nu}(l) = \bar{\nu}(p)$ for all $l \rightsquigarrow p \in \mathcal{P}_2$. Then \mathcal{P} is chain-free if and only if \mathcal{P}_2 is.*

Proof. If \mathcal{P} is chain-free then obviously \mathcal{P}_2 is chain-free. For the other direction, suppose \mathcal{P}_2 is chain-free and that we have a minimal dependency chain $[(\rho_i, s_i, t_i) \mid i \in \mathbb{N}]$ over \mathcal{P} and suitable projection function ν . As before, we can assume this chain does not use **beta**, so always $\rho_i = l_i \rightsquigarrow p_i$. As the subterm relation is stable, $\bar{\nu}(l_i) \triangleright \bar{\nu}(p_i)$ implies that $\bar{\nu}(s_i) \triangleright \bar{\nu}(t_i)$.

Since each $t_i \rightarrow_{\mathcal{R}}^* s_{i+1}$, we have a reduction $\bar{\nu}(s_1) \triangleright \bar{\nu}(t_1) \rightarrow_{\mathcal{R}}^* \bar{\nu}(s_2) \triangleright \bar{\nu}(t_2) \rightarrow_{\mathcal{R}}^* \dots$. Since \mathcal{P}_2 is chain-free, $\rho_i \in \mathcal{P}_i$ for infinitely many i , so $\bar{\nu}(s_i) \triangleright \bar{\nu}(t_i)$ infinitely often. Thus we have an infinite $\rightarrow_{\mathcal{R}} \cup \triangleright$ reduction starting in a terminating term $\bar{\nu}(s_1)$, contradiction. \square

Example 8.3. Both of the remaining SCCs of Example 7.9 are chain-free, as is evident with a projection $\nu(\text{lreq}^\sharp) = 1$ for the SCC $\{(1)\}$ and $\nu(\text{from}^\sharp) = 2$ for the SCC $\{(5)\}$.

8.2. Usable Rules. Formative rules provide a nice counterpart of usable rules, but it would be even nicer if we had both! Usable rules can be defined, but with severe restrictions; we follow the ideas in [36], where usable rules are defined for static dependency pairs.

Definition 8.4 (Usable Rules). A term s is considered *risky* if it has a subterm $x \cdot t$ with $x \in FV(s)$. Let $f \sqsubseteq_{us} g$ if there is a rewrite rule $f(l_1, \dots, l_n) \rightarrow r$ where either r contains the symbol g , or r is risky, or r is an abstraction or variable of functional type. The reflexive-transitive closure of \sqsubseteq_{us} is denoted by \sqsubseteq_{us}^* . Overloading notation, let $s \sqsubseteq_{us}^* g$ denote that either s contains a symbol f with $f \sqsubseteq_{us}^* g$, or s is risky.

The set of usable rules of a term s , notation $UR(s, \mathcal{R})$, consists of those rules $f(l_1, \dots, l_n) \cdot l_{n+1} \dots l_m \rightarrow r \in \mathcal{R}$ such that $s \sqsubseteq_{us}^* f$. The set of usable rules of a non-collapsing dependency pair $l \rightsquigarrow f(p_1, \dots, p_n) \cdot p_{n+1} \dots p_m$ is the union $UR(p_1, \mathcal{R}) \cup \dots \cup UR(p_m, \mathcal{R})$. The set of usable rules of a set of dependency pairs, $UR(\mathcal{P}, \mathcal{R})$, is the union $\bigcup_{\rho \in \mathcal{P}} UR(\rho, \mathcal{R})$ if \mathcal{P} is non-collapsing and just \mathcal{R} otherwise.

Example 8.5. Consider an AFS for list manipulation which has four rules:

$$\begin{aligned} \text{map}(F, \text{nil}) &\rightarrow \text{nil} \\ \text{map}(F, \text{cons}(h, t)) &\rightarrow \text{cons}(F \cdot h, \text{map}(F, t)) \\ \text{append}(\text{nil}, l) &\rightarrow l \\ \text{append}(\text{cons}(h, t), l) &\rightarrow \text{cons}(\text{append}(h, t), l) \end{aligned}$$

The dependency graph of this system has two SCCs: the set $\{\text{map}^\sharp(F, \text{cons}(h, t)) \rightsquigarrow F \cdot h, \text{map}^\sharp(F, \text{cons}(h, t)) \rightsquigarrow \text{map}^\sharp(F, t)\}$ and the set $\{\text{append}^\sharp(\text{cons}(h, t), l) \rightsquigarrow \text{append}^\sharp(h, t)\}$. The former contains a dependency pair with $F \cdot h$ in the right-hand side, so its usable rules are just \mathcal{R} . But the latter set has only the **append** rules as usable rules.

Very similar to the first-order case, we can prove the following result:

Theorem 8.6. Let $\mathcal{P} = \mathcal{P}_1 \uplus \mathcal{P}_2$ be a set of non-collapsing dependency pairs, and suppose there is a reduction pair (\succeq, \succ) such that:

- $\bar{l} \succ \bar{p}$ for all $l \rightsquigarrow p \in \mathcal{P}_1$
- $\bar{l} \succeq \bar{p}$ for all $l \rightsquigarrow p \in \mathcal{P}_2$
- $l \succeq r$ if $l \rightarrow r \in UR(\mathcal{P}, FR(\mathcal{P}))$
- $\mathbf{p}_\sigma(x, y) \succeq x, y$ for fresh symbols $\mathbf{p}_\sigma : [\sigma \times \sigma] \Rightarrow \sigma$.

Then \mathcal{P} is chain-free if and only if \mathcal{P}_2 is chain-free.

Here, the pair \bar{l}, \bar{p} is determined from l, p in a systematic way like in Section 5.5. The $\mathbf{p}_\sigma(x, y) \succeq x, y$ constraints are trivially oriented both with algebra interpretations and CPO.

The proof of Theorem 8.6 takes some work, but has no novelties compared to the proof for the static method in [4]. We shall be relatively brief about it.

Proof Sketch of Theorem 8.6. Let $\mathbf{p}_\sigma : [\sigma \times \sigma] \Rightarrow \sigma$ and $\perp_\sigma : \sigma$ be new symbols for all types σ , and let \mathcal{C}_ϵ be the set of all rules $\mathbf{p}_\sigma(x, y) \rightarrow x, \mathbf{p}_\sigma(x, y) \rightarrow y$. We will see that any minimal dependency chain $[(\rho_i, s_i, t_i) \mid i \in \mathbb{N}]$ with all $\rho_i \in \mathcal{P}$ can be transformed into a (not necessarily minimal) dependency chain $[(\rho_i, s'_i, t'_i) \mid i \in \mathbb{N}]$ which uses the same dependency pairs, but where $t'_i \xrightarrow{*}_{UR(\mathcal{P}, FR(\mathcal{P})) \cup \mathcal{C}_\epsilon} s'_{i+1}$ for all i . Following the proof of Theorems 5.16 and 5.21, if a reduction pair as described in the theorem exists, then this transformed chain must use dependency pairs in \mathcal{P}_2 infinitely often, so the same holds for the original. Thus, if \mathcal{P} is not chain-free, then neither is \mathcal{P}_2 . Of course, if \mathcal{P} is chain-free then so is its subset \mathcal{P}_2 .

To transform the minimal dependency chain $[(\rho_i, s_i, t_i) \mid i \in \mathbb{N}]$, we note that:

(I) *every term has only finitely many direct reducts.*

This is obvious because \mathcal{R} is finite. Let $\mathcal{R}_1 := FR(\mathcal{P})$ and say a symbol f is a *usable symbol* if $p \sqsupseteq_{us}^* f$ for some dependency pair $l \rightsquigarrow p \in \mathcal{P}$ (where \sqsupseteq_{us}^* is based on \mathcal{R}_1 rather than \mathcal{R}), or f is a constructor symbol. We assume:

(II) $UR(\mathcal{P}, \mathcal{R}_1) \neq \mathcal{R}_1$.

This is a safe assumption, because if $UR(\mathcal{P}, \mathcal{R}_1) = \mathcal{R}_1$ the theorem is automatically satisfied. Now, because of (I) we can define the function φ on terminating terms, given by the clauses:

- $\varphi(\lambda x. u) = \lambda x. \varphi(u)$
- $\varphi(x \cdot u_1 \cdots u_n) = x \cdot \varphi(u_1) \cdots \varphi(u_n)$
- $\varphi(f(u_1, \dots, u_n) \cdot u_{n+1} \cdots u_m) = f(\varphi(u_1), \dots, \varphi(u_n)) \cdot \varphi(u_{n+1}) \cdots \varphi(u_m)$ if f is a usable symbol
- $\varphi(s) = p_\sigma(f(\varphi(u_1), \dots, \varphi(u_n)) \cdot \varphi(u_{n+1}) \cdots \varphi(u_m), D_\sigma(\{t \mid s \rightarrow_{\mathcal{R}_1} t\}))$ if $s : \sigma$ and $s = f(u_1, \dots, u_n) \cdot u_{n+1} \cdots u_m$ with f not a usable symbol
- $\varphi(s) = D_\sigma(\{t \mid s \rightarrow_{\mathcal{R}_1} t\})$ if $s : \sigma$ and s does not have any of these forms

Here, $D_\sigma(\emptyset) = \perp_\sigma$ and $D_\sigma(X) = p_\sigma(\varphi(u), D_\sigma(X \setminus \{u\}))$ if X is non-empty and u is lexicographically its smallest element (this is only defined for finite sets X).

Writing γ^φ for a substitution $[x := \varphi(\gamma(x)) \mid x \in \text{dom}(\gamma)]$, it is not hard to see that:

(III) *if s has no subterms $(\lambda x. t) \cdot u$ or $x \cdot u$ with $x \in \text{dom}(\gamma)$, then $\varphi(s\gamma) \rightarrow_{\mathcal{C}_e}^* s\gamma^\varphi$. If all symbols occurring in s are usable symbols, then even $\varphi(s\gamma) = s\gamma^\varphi$.*

The proof holds by induction on the form of s , noting that the last case of the definition of φ is never applicable. With induction on the size of X it follows easily that $D_\sigma(X) \rightarrow_{\mathcal{C}_e}^* \varphi(s)$ for any $s \in X$. Combining this with (III) we can derive:

(IV) *if $s \rightarrow_{\mathcal{R}_1} t$ with s terminating, then $\varphi(s) \rightarrow_{UR(\mathcal{P}, \mathcal{R}_1) \cup \mathcal{C}_e}^* \varphi(t)$.*

This holds with induction on the form of s ; all the induction cases are trivial either with the induction hypothesis or with the observation that $D_\sigma(X) \rightarrow_{\mathcal{C}_e}^* \varphi(u)$ if $u \in X$. By the same observation, the base case (a headmost step) is easy if $s = f(\vec{u}) \cdot \vec{v}$ with f not a usable symbol, or if s reduces with a headmost β -step. What remains is the case when $s = f(\vec{u}) \cdot \vec{v}$ with f a usable symbol, and the reduction is headmost: $s = l\gamma \cdot v_{k+1} \cdots v_m$ and $t = r\gamma \cdot v_{k+1} \cdots v_m$. But since f is a usable symbol, $l \rightarrow r$ is a usable rule, so since rules are β -normal: $\varphi(s) = \varphi(l\gamma) \cdot \varphi(v_{k+1}) \cdots \varphi(v_m) \rightarrow_{\mathcal{C}_e}^* l\gamma^\varphi \cdot \varphi(v_{k+1}) \cdots \varphi(v_m) \rightarrow_{UR(\mathcal{P}, \mathcal{R}_1)} r\gamma^\varphi \cdot \varphi(v_{k+1}) \cdots \varphi(v_m) = \varphi(r\gamma) \cdot \varphi(v_{k+1}) \cdots \varphi(v_m)$ by (III), which equals $\varphi(t)$ because either r has the form $g(\vec{w}) \cdot \vec{q}$, or $m = k$ (since otherwise $UR(\mathcal{P}, \mathcal{R}_1) = \mathcal{R}_1$, contradicting (II)).

Now we are almost there. Let $[(\rho_i, s_i, t_i) \mid i \in \mathbb{N}]$ be a minimal dependency chain with all $\rho_i \in \mathcal{P}$, so for each i we can write $\rho_i = l_i \rightsquigarrow p_i$ and $s_i = l_i \gamma_i$, $t_i = p_i \gamma_i$. Let $s'_i = l_i \gamma_i^\varphi$ and $t'_i = p_i \gamma_i^\varphi$ (this is well-defined because the strict subterms of s_i and t_i are terminating). Then by (III) and (IV) $p_i \gamma_i^\varphi \rightarrow_{UR(\mathcal{P}, FR(\mathcal{P})) \cup \mathcal{C}_e, in}^* l_{i+1} \gamma_{i+1}^\varphi$ as required. \square

The notion of usable rules is in particular useful for systems with a first-order subset, as the usable rules of first-order dependency pairs will typically be first-order. As discussed in [8], we can use this to apply first-order termination proving techniques (and termination tools!) to a part of a higher-order dependency pair problem.

9. ALGORITHM

In this section, we combine all the results so far in a ready-to-use algorithm.

All the results in this paper can be readily combined in one algorithm, to prove termination of an AFS (local or not):

- (1) • determine whether the system is local (Definition 6.1),
 - complete the system (Definition 5.1),
 - determine its dependency pairs (Definition 5.4)
 - calculate an approximation G for the dependency graph (Section 5.3);
- (2) remove all nodes and edges from G which are not part of a cycle (Section 5.3);
- (3) if G is empty return **terminating**; otherwise choose an SCC \mathcal{P} (Section 5.3);
- (4) if \mathcal{P} is non-collapsing, find a projection function ν such that $\bar{\nu}(l) \supseteq \bar{\nu}(p)$ for all $l \rightsquigarrow p \in \mathcal{P}$, and $\bar{\nu}(l) \triangleright \bar{\nu}(p)$ at least once (Section 8.1); if this succeeds, remove all strictly oriented pairs from G and continue with 2, otherwise continue with 5;
- (5) a. if \mathcal{P} is non-collapsing:
 - let $S := \emptyset$ and $\Sigma := \mathcal{F}_c^\sharp$;
 - let $\psi(s)$ be defined as just s ;
 - let $A := UR(\mathcal{P}, FR(\mathcal{P}))$ if \mathcal{R} is local, $A := UR(\mathcal{P}, \mathcal{R})$ otherwise (Definition 6.10, 8.4);
 b. if \mathcal{P} is collapsing and \mathcal{R} is not local:
 - let $S := \mathcal{F}$ and $\Sigma := \mathcal{F}_c^\sharp$;
 - let $\psi(s)$ be defined as just s ;
 - let $A := \mathcal{R} \cup \{f(\vec{x}) \rightarrow f^\sharp(\vec{x}) \mid f \in \mathcal{D}\}$;
 c. if \mathcal{P} is collapsing and \mathcal{R} is local:
 - let S be the set of function symbols $f^- : \sigma$ where $f : \sigma \in \mathcal{F}$ and f occurs below an abstraction in a right-hand side of $FR(\mathcal{P}) \cup \mathcal{P}$, and let $\Sigma := \mathcal{F}_c^\sharp \cup S$;
 - let $\psi(s)$ be defined as $\text{tag}(s)$ (Definition 6.3);
 - let $A := FR(\mathcal{P})^{\text{tag}} \cup \{f^-(\vec{x}) \rightarrow f(\vec{x}), g^-(\vec{x}) \rightarrow g^\sharp(\vec{x}) \mid f^-, g^- \in S, g \in \mathcal{D}\}$;
- (6) determine a partitioning $\mathcal{P} = \mathcal{P}_1 \uplus \mathcal{P}_2$ and a reduction pair (Section 5.4) (\succeq, \succ) with:
 - (a) $l \cdot x_1 \cdots x_n \succ \psi(p) \cdot c_{\sigma_1} \cdots c_{\sigma_m}$ for $l \rightsquigarrow p \in \mathcal{P}_1$ (both sides base type, fresh \vec{x});
 - (b) $l \cdot x_1 \cdots x_n \succeq \psi(p) \cdot c_{\sigma_1} \cdots c_{\sigma_m}$ for $l \rightsquigarrow p \in \mathcal{P}_2$ (both sides base type, fresh \vec{x});
 - (c) $l \succeq r$ for $l \rightarrow r \in A$;
 - (d) either \mathcal{P} is non-collapsing, or (\succeq, \succ) respects \triangleright^S (Definition 6.19);
 - (e) if \mathcal{R} is local and \mathcal{P} collapsing:
 - for all $f \in \mathcal{F}^- \setminus S$: $f^-(\vec{x}) \succeq f(\vec{x}), f^\sharp(\vec{x})$ and $f^-(\vec{x}) \cdot \vec{y} \succeq x_i \cdot \vec{c}$ (**).

if this step fails, return **fail**;

suitable reduction pairs can be found with e.g. weakly monotonic algebras (Section 7.1), or argument functions and CPO (Section 7.2)

- (7) remove all pairs in \mathcal{P}_1 from the graph, and continue with (2).

(**) Since the symbols $f \in \mathcal{F}^- \setminus S$ do not occur in any of the other constraints, all the common reduction pairs (CPO, weakly monotonic algebras and most existing first-order techniques) can easily be extended to satisfy these requirements; for example, given a path ordering which satisfies the rest of the requirements, just add “ $f^- >_F f$ ” to the function symbol precedence. That is why we split off these requirements rather than including the symbols in S , and why we generally ignore this restriction on (\succeq, \succ) in examples.

This algorithm iterates over the graph approximation, removing nodes and edges until none remain; this technique originates in [14]. It is justified by the following observation:

Lemma 9.1. *Let G be an approximation of the dependency graph of an AFS \mathcal{R} , and let \mathcal{P} be a set of dependency pairs. Suppose that every SCC in the subgraph of G which contains only nodes in \mathcal{P} is chain-free. Then also \mathcal{P} is chain-free.*

Proof. A trivial adaptation of the proof of Lemma 5.12. \square

The algorithm seeks to prove that $\text{DP}(\mathcal{R})$ is either tagged-chain-free (if \mathcal{R} is local) or chain-free (if not), which suffices by Theorems 6.15 and 5.7 respectively.

- (1) To prove that $\text{DP}(\mathcal{R})$ is (tagged-)chain-free, it suffices to prove that all SCCs of G are either tagged-chain-free (if \mathcal{R} is local and \mathcal{P} is non-collapsing) or chain-free (if \mathcal{R} is not local or \mathcal{P} is collapsing); this is valid by Lemmas 5.12 and 8.1.
- (2) Thus, until G is empty, we choose an SCC \mathcal{P} and show that \mathcal{P} is tagged-chain-free or chain-free as required if some strict subset \mathcal{P}' is (tagged-)chain-free (using the subterm criterion or a reduction pair, which is justified by Theorems 5.21, 6.20 and 8.2). By Lemma 9.1 this is the case if all sub-SCCs $\mathcal{P}_1, \dots, \mathcal{P}_n$ of \mathcal{P}' are (tagged-)chain-free. To see that all remaining SCCs of G and $\mathcal{P}_1, \dots, \mathcal{P}_n$ are (tagged-)chain-free, we remove all strictly oriented pairs from G , and remove all nodes not on a cycle, so the SCCs of the result are exactly the sets that must be proved (tagged-)chain-free.

Example 9.2. We apply the algorithm on the AFS `eval`, with the following symbols:

$$\begin{array}{lll} \text{o} : \text{M} & \text{dom} : [\text{M} \times \text{M} \times \text{M}] \Rightarrow \text{M} & \text{eval} : [\text{M} \times \text{M}] \Rightarrow \text{M} \\ \text{s} : [\text{M}] \Rightarrow \text{M} & \text{fun} : [\text{M} \Rightarrow \text{M} \times \text{M} \times \text{M}] \Rightarrow \text{M} & \end{array}$$

and the following rewrite rules:

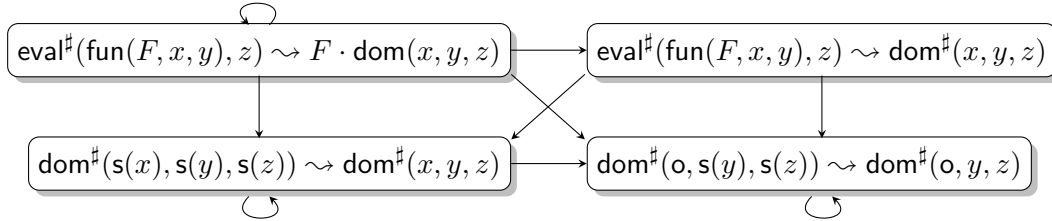
$$\begin{array}{lll} \text{dom}(\text{s}(x), \text{s}(y), \text{s}(z)) & \rightarrow & \text{s}(\text{dom}(x, y, z)) & \text{dom}(x, y, \text{o}) & \rightarrow & x \\ \text{dom}(\text{o}, \text{s}(y), \text{s}(z)) & \rightarrow & \text{s}(\text{dom}(\text{o}, y, z)) & \text{dom}(\text{o}, \text{o}, z) & \rightarrow & \text{o} \\ \text{eval}(\text{fun}(F, x, y), z) & \rightarrow & F \cdot (\text{dom}(x, y, z)) & & & \end{array}$$

The `fun` symbol represents a function over the natural numbers with an interval it is defined on; `eval` calculates its value in a point, provided the point is in the domain of the function. We prove termination of `eval` by following the algorithm.

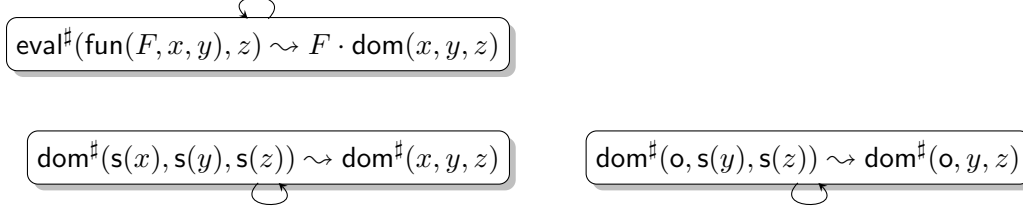
(1): *complete the rules, calculate the dependency pairs and approximate the dependency graph.* The AFS `eval` is local. Because `eval` has no rules of the form $l \rightarrow \lambda x. r$, the completed system is the same as the original system. We have the following four dependency pairs:

$$\begin{array}{ll} \text{eval}^\#(\text{fun}(F, x, y), z) \rightsquigarrow F \cdot \text{dom}(x, y, z) & \text{dom}^\#(\text{s}(x), \text{s}(y), \text{s}(z)) \rightsquigarrow \text{dom}^\#(x, y, z) \\ \text{eval}^\#(\text{fun}(F, x, y), z) \rightsquigarrow \text{dom}^\#(x, y, z) & \text{dom}^\#(\text{o}, \text{s}(y), \text{s}(z)) \rightsquigarrow \text{dom}^\#(\text{o}, y, z) \end{array}$$

We will use the following (approximation of the) dependency graph:



Step 2: remove nodes and edges not part of a cycle. This removes the second eval^\sharp dependency pair, and leaves the following graph:



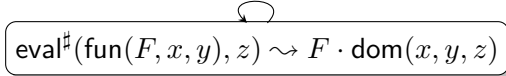
*Step 3: return **terminating** if G is empty, otherwise choose an SCC \mathcal{P} .* Since the graph is not empty, we choose $\mathcal{P} = \{\text{dom}^\sharp(o, s(y), s(z)) \rightsquigarrow \text{dom}^\sharp(o, y, z)\}$.

Step 4: apply the subterm criterion if possible. The chosen set is indeed non-collapsing; we choose a projection function $\nu(\text{dom}^\sharp) = 2$, and have $\bar{\nu}(\text{dom}^\sharp(o, s(y), s(z))) = s(y) \triangleright u = \bar{\nu}(\text{dom}^\sharp(o, y, z))$, and can remove this dependency pair from G .



Step 2: remove nodes and edges not part of a cycle. Everything is on a cycle.

Step 3,4: choose the next SCC and apply the subterm criterion. We choose the set $\mathcal{P} = \{\text{dom}^\sharp(s(x), s(y), s(z)) \rightsquigarrow \text{dom}^\sharp(x, y, z)\}$, which is non-collapsing, and apply the subterm criterion with $\nu = 2$. Since $s(y) \triangleright y$, this leaves the graph with a single node:



Step 2: remove nodes and edges not part of a cycle. Everything is on a cycle.

*Step 3: return **terminating** if G is empty, otherwise choose an SCC \mathcal{P} .* Since the graph is not empty, we choose $\mathcal{P} = \{\text{eval}^\sharp(\text{fun}(F, x, y), z) \rightsquigarrow F \cdot \text{dom}(x, y, z)\}$.

Step 4: apply the subterm criterion if possible. This is not possible, as \mathcal{P} is collapsing.

Step 5: determine Σ , S , A and ψ . Noting that the system is local, but has abstractions in neither left- nor right-hand sides, we let $\Sigma := \{o, s, \text{dom}, \text{fun}, \text{eval}, \text{dom}^\sharp, \text{eval}^\sharp\}$ and $S := \emptyset$, while ψ is the tagging function. Since $S = \emptyset$, we calculate $A = FR(\mathcal{P}) = FR(\text{eval}^\sharp(\text{fun}(F, x, y), z) \rightsquigarrow F \cdot \text{dom}(x, y, z)) = FR(\text{fun}(F, x, y)) \cup FR(z) = FR(\text{fun}(F, x, y))$, which consists of the rules $\text{eval}(\text{fun}(F, x, y), z) \rightarrow F \cdot \text{dom}(x, y, z)$, $\text{dom}(x, y, o) \rightarrow x$ and $\text{dom}(o, o, z) \rightarrow o$ (since not $\text{fun} \sqsubseteq_{fo}^* s$).

Step 6: determine a suitable partitioning and reduction pair. As \mathcal{P} consists of only one dependency pair, we must choose $\mathcal{P}_1 = \mathcal{P}$. We have the following ordering constraints:

$$\begin{array}{lll} \text{eval}^\sharp(\text{fun}(F, x, y), z) >_F F \cdot \text{dom}(x, y, z) & \text{dom}(x, y, o) \succeq x & \\ \text{eval}(\text{fun}(F, x, y), z) \succeq F \cdot \text{dom}(x, y, z) & \text{dom}(o, o, z) \succeq o & \end{array}$$

The first constraint is for the dependency pair, the others for the formative rules. Using an argument function π which maps $\text{dom}(x, y, z)$ to $\text{dom}'(x, y)$ and otherwise maps $f(\vec{x})$ to itself (π is an *argument filtering*), this leaves the following constraints:

$$\begin{array}{lll} \text{eval}^\sharp(\text{fun}(F, x, y), z) >_F F \cdot \text{dom}'(x, y) & \text{dom}'(x, y) \succeq x & \\ \text{eval}(\text{fun}(F, x, y), z) \succeq F \cdot \text{dom}'(x, y) & \text{dom}'(o, o) \succeq o & \end{array}$$

This is satisfied with $(\succeq_{\text{CPO}}, >_{\text{CPO}})$ from Section 7.2, using a precedence $\text{fun} >_F \text{dom}' >_F s, o$.

Steps 7,2,3: remove all pairs in \mathcal{P}_1 and everything that's not on a cycle anymore. We remove the last node; the graph is empty. We conclude: this AFS is terminating.

10. EXTENSIONS

In this section we discuss two improvements to the method in this paper: a combination with the static dependency pair approach, and a way to deal with AFSs with infinitely many rules, in particular those generated from a polymorphic AFS.

10.1. Merging the Static and Dynamic Approach. The dynamic style of dependency pairs presented in this paper is not the only way to do dependency pairs; as discussed in Section 2 there are several strong results in *static* dependency pairs. Most importantly, in a static dependency pair approach we do not have to consider collapsing dependency pairs, and hence the dependency graph is usually simpler. Usable rules and the subterm criterion are more often applicable, and argument filterings can be used without restrictions.

Since the dynamic approach is applicable to a larger class of AFSs, but the static approach gives easier constraints, it seems sensible for a termination tool to implement both. Static dependency pairs are defined for HRSs, but extending the proof to AFSs takes next to no effort, and using Lemma 6.13 we can add formative rules to it as well. Below, we give a short overview without proofs; for the full work on static dependency pairs, see [26, 36].

Definition 10.1 (AFS with Base Output Types). An AFS $(\mathcal{F}, \mathcal{R})$ has *base output types* if for all $f : [\sigma_1 \times \dots \times \sigma_n] \Rightarrow \tau \in \mathcal{F}$, the output type τ is a base type.

The AFS for `eval` from Example 9.2 has base output types, but the AFS `twice` does not, because of the symbol `twice` : $[\text{nat} \Rightarrow \text{nat}] \Rightarrow \text{nat} \Rightarrow \text{nat}$. It is always possible to transform an AFS into an AFS with base output types by η -expanding the rules [21], without losing non-termination; however, termination may be lost by this transformation.

Definition 10.2 (Plain Function Passing AFS). An AFS is *plain function passing* (PFP) if for all rules $f(l_1, \dots, l_n) \rightarrow r$: if r contains a functional variable F , then F is one of the l_i .

The notion is slightly simplified from the original definition, for the sake of easy explanation.

The set of *static dependency pairs* of a PFP AFS with base output types consists of the pairs $l^\# \rightsquigarrow g^\#(\vec{s})$ with $l \rightarrow r$ a rewrite rule, $r \supseteq g(\vec{s})$, and g a defined symbol. Note that the right-hand side of a static dependency pair may contain variables which do not occur in its left-hand side. For example, the rewrite rule $\text{I}(\text{s}(n)) \rightarrow \text{twice}(\lambda x. \text{I}(x), n)$ has two static dependency pairs: $\text{I}^\#(\text{s}(n)) \rightsquigarrow \text{twice}^\#(\lambda x. \text{I}(x), n)$ and $\text{I}^\#(\text{s}(n)) \rightsquigarrow \text{I}^\#(x)$.

A *static dependency chain* is a sequence $[(\rho_i, s_i, t_i) \mid i \in \mathbb{N}]$ where:

- each ρ_i is a static dependency pair $l \rightsquigarrow p$;
- there is some substitution γ such that $s_i = l\gamma$ and $t_i = r\gamma$;
- each $t_i \rightarrow_{FR(\rho_{i+1})}^* s_{i+1}$ (where $FR(\rho_{i+1}) = \mathcal{R}$ in a non-local AFS);
- the immediate subterms of all t_i are “computable” (which implies termination).

Claim 10.3. *A plain function passing AFS with base output types is terminating if it does not have a static dependency chain.*

Claim 10.3 can be verified by mimicking the proof in [26] for the different setting, and additionally using Lemma 6.13; this proof contains no novelties.

Observing that all static dependency pairs are non-collapsing, it is worth noting that a static dependency chain is also a dynamic minimal dependency chain as defined in Definition 5.6; just for a different set of dependency pairs. In the results of this paper, we did not use the property that the right-hand side of dependency pairs contains no new variables; thus, we can also use the algorithm from Section 9 to prove absence of static dependency chains. With static dependency pairs, we can always use case 5a, so we have unrestricted argument filterings, the subterm criterion and usable rules, as we had in the HRS setting by [36].

Now an automatic tool could, for example, try dynamic dependency pairs, and if that fails but the AFS is PFP, η -expand the rules and use the same module to attempt a static approach. But in some cases we can do better. For instance, the dynamic dependency pairs for `map` from Example 7.6 are:

$$\begin{aligned} (A) \quad \text{map}^\#(F, \text{cons}(h, t)) &\rightsquigarrow F \cdot h \\ (B) \quad \text{map}^\#(F, \text{cons}(h, t)) &\rightsquigarrow \text{map}^\#(F, t) \end{aligned}$$

In the static approach, we only have pair (B). Thus, if there is a dependency chain on the static dependency pairs, there is also one on the dynamic ones. In cases like this, there is no point using the dynamic approach instead of the static one.

Definition 10.4. We say an AFS $(\mathcal{F}, \mathcal{R})$ is *strongly plain function passing* (SPFP) if it is plain function passing, has base output types, and the right-hand sides of \mathcal{R} do not have subterms of the form $\lambda x. C[f(s_1, \dots, s_n)]$ where f is a defined symbol and $x \in FV(f(\vec{s}))$.

Theorem 10.5. *A SPFP AFS admits a minimal dynamic dependency chain if it admits a minimal static dependency chain. A SPFP and left-linear AFS admits a minimal static dependency chain if and only if it is non-terminating.*

Proof. Given a SPFP AFS, all its static dependency pairs are also dynamic dependency pairs, so any minimal static dependency chain is a minimal dynamic dependency chain.

If the SPFP AFS is moreover left-linear, then termination implies that there is no minimal dynamic dependency chain (by Theorem 5.18), let alone a static one. \square

This theorem gives a strengthening of the dynamic approach, by allowing us to remove collapsing dependency pairs in an AFS if it is strongly plain function passing. It also provides a completeness result for the static approach, for the class of SPFP systems.

In summary, we can combine the static and dynamic approaches as follows:

```

if  $\mathcal{R}$  is strongly plain function passing: return DPframework(STATIC)
if DPframework(DYNAMIC) = TERMINATING: return TERMINATING
if  $\mathcal{F}$  does not have base output types:  $\eta$ -expand  $(\mathcal{F}, \mathcal{R})$ 
return DPframework(STATIC)

```

Here, `DPframework` is the algorithm of Section 9, and the argument (`STATIC` or `DYNAMIC`) determines how the dependency pairs are calculated. Alternatively, the two frameworks may be run in parallel, or we might consider optimisations to avoid double work. At present, however, we have not done this, because testing on the current termination problem database shows no improvement in strength obtained from using the static approach after the dynamic approach has failed (although using the static approach for SPFP AFSs *does* significantly improve the power). Experimental results and statistics are given in Section 11.

10.2. Polymorphism. In this paper we consider monomorphic AFSs. In other definitions of the AFS formalism, a kind of polymorphism à la ML is admitted. In a polymorphic AFS, we can have for instance a function symbol $\text{if} : [\text{bool} \times \alpha \times \alpha] \Rightarrow \alpha$ with α a type variable, and rewrite rules $\text{if}(\text{true}, x, y) \rightarrow x$ and $\text{if}(\text{false}, x, y) \rightarrow y$.

Here α can be instantiated by base types, but also by functional types such as $\text{nat} \Rightarrow \text{nat}$. A polymorphic AFS can be transformed into a monomorphic AFS with infinitely many rewrite rules by considering all possible instantiations for the type variables in the rewrite rules. In the example above, we get for instance $\text{if}_{[\text{bool} \times \text{nat} \times \text{nat}] \Rightarrow \text{nat}}(\text{true}, x, y) \rightarrow x$ and $\text{if}_{[\text{bool} \times (\text{nat} \Rightarrow \text{nat}) \times (\text{nat} \Rightarrow \text{nat})] \Rightarrow \text{nat} \Rightarrow \text{nat}}(\text{false}, x, y) \rightarrow y$. Such a transformation is discussed in [21].

In this paper we considered monomorphic AFSs with finitely many rules, so we can neither deal with polymorphic AFSs directly nor with their transformations to monomorphic AFSs. However, a further inspection of the proofs shows that finiteness of the set of rewrite rules is primarily used in dealing with the dependency graph. So suppose we did not pose this restriction, and went on as before. This leads to a dependency graph with infinitely many nodes. Now consider the following definition of a dependency graph approximation:

Definition 10.6 (Infinite Dependency Graph Approximation). An (infinite) dependency graph A is approximated by the finite graph G if there is a mapping f from the nodes in A to the nodes in G , such that: if A contains an edge from node a to node b , then G contains an edge from node $f(a)$ to node $f(b)$.

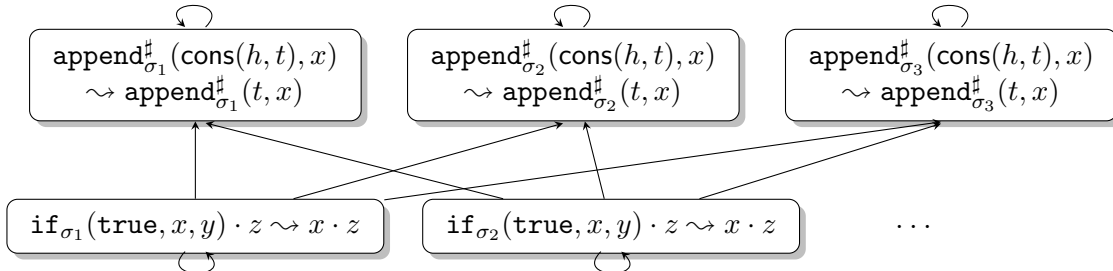
The original notion of a dependency graph approximation is an instance of this definition, where f is the identity. With this definition, all proofs go through essentially unmodified, only when considering a set \mathcal{P} of nodes, we are interested not in the elements of \mathcal{P} but in the corresponding dependency pairs. For instance in Theorem 5.16, we must see that $l \succ p$ for all $l \rightsquigarrow p \in \text{DP}$ with $f(l \rightsquigarrow p) \in \mathcal{P}_1$ and $l \succeq p$ for all $l \rightsquigarrow p \in \text{DP}$ with $f(l \rightsquigarrow p) \in \mathcal{P}_2$.

Example 10.7. Consider the polymorphic AFS with if as defined before, and $\text{append} : [\text{list}(\alpha) \times \text{list}(\alpha)] \Rightarrow \text{list}(\alpha)$.

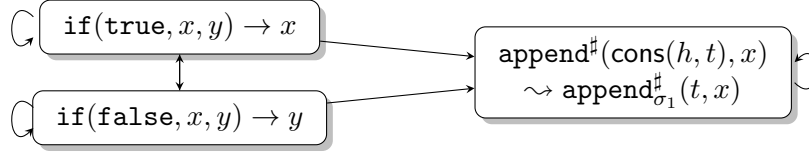
$$\begin{aligned} \text{if}(\text{true}, x, y) &\rightarrow x \\ \text{if}(\text{false}, x, y) &\rightarrow y \\ \text{append}(\text{nil}, x) &\rightarrow x \\ \text{append}(\text{cons}(h, t), x) &\rightarrow \text{cons}(h, \text{append}(t, x)) \end{aligned}$$

The corresponding monomorphic AFS has infinitely many rules, and infinitely many dependency pairs. All the rules generated from the append rules (such as append for $\text{list}(\text{nat})$ or for $\text{list}(\text{nat} \Rightarrow \text{bool})$) have only one dependency pair $\text{append}_{\sigma}^{\#}(\text{cons}(h, t), x) \rightsquigarrow \text{append}_{\sigma}^{\#}(t, x)$. The if rules of basic types (such as when α is instantiated with nat) have no dependency pairs, but if we for instance consider $\sigma := [\text{bool} \times \text{nat} \Rightarrow \text{nat} \times \text{nat} \Rightarrow \text{nat}] \Rightarrow \text{nat} \Rightarrow \text{nat}$ then the rule $\text{if}_{\sigma}(\text{true}, x, y) \rightarrow x$ does give a dependency pair $\text{if}_{\sigma}(\text{true}, x, y) \cdot z \rightsquigarrow x \cdot z$.

The dependency graph is infinite, and part of it could be sketched like this:



For a dependency graph approximation, it makes sense to combine all the “similar” dependency pairs in one node. We choose the following finite graph approximation:



Here, all dependency pairs $\text{append}^\#_\sigma(\text{cons}(h, t), x) \rightsquigarrow \text{append}^\#_\sigma(t, x)$ are mapped to the **append** node, and the $\text{if}_\sigma(\text{true}, x, y) \cdot z_1 \cdots z_n \rightsquigarrow x \cdot z_1 \cdots z_n$ are mapped to $\text{if}(\text{true}, x, y) \rightarrow x$. Note the different arrows: the **append** node uses a \rightsquigarrow arrow, to indicate that it catches dependency pairs of the *first kind* (see Definition 5.4) which have the given form, whereas the **if** nodes use a \rightarrow arrow and catch dependency pairs of the *second kind*.

Example 10.7 gives a good suggestion for how to split the dependency pairs into a finite number of groups. For a polymorphic rule, we generate the set of dependency pairs of the first kind as before, $\text{DP}_1(\mathcal{R}) = \{l^\# \rightsquigarrow p^\# \mid l \rightarrow r \in \mathcal{R} \mid p \in \text{Cand}(r)\}$; for all instances of the rule, the dependency pairs of the first kind will be type instantiations of a pair in $\text{DP}_1(\mathcal{R})$. For the nodes of the graph approximation we choose $\text{DP}_1(\mathcal{R})$, and in addition those rules $l \rightarrow r$ whose type is not a base type and where $\text{head}(r)$ is either a variable or a term $f(\vec{r})$ with $f \in \mathcal{D}$. Now each dependency pair in the instantiated system is either mapped to the corresponding pair in $\text{DP}_1(\mathcal{R})$, or, for a pair of the second kind, to the rule that generated it.

With this choice for the graph approximation we can avoid any infinite reasoning, *if* we can find a reduction pair that is equipped to prove statements of the form “ $\bar{l} \succ \bar{p}$ for all type instantiations” given l and p . Finding such reduction pairs is the object of separate study, however. As for formative rules, it is not hard to devise an algorithm which, given a symbol s , finds a set of polymorphic rules PFR such that the rewrite relation generated by PFR is included in $\rightarrow_{FR(s)}$ (the difficulty is in determining how good the result of such an algorithm is, as it may well include more rules than strictly necessary). Usable rules require only that every term has only finitely many reducts (item I in the proof of Theorem 8.6); this is the case for systems generated from a finite set of polymorphic rules as well.

11. EXPERIMENTAL RESULTS

In this section, we present the performance of the higher-order termination tool WANDA in different settings: using only dynamic dependency pairs, using only static dependency pairs, using both dynamic and static dependency pairs, and using only rule removal.

The dependency pair framework described in this paper, and in particular the algorithm of Section 9, forms the core of the higher-order termination tool WANDA [20]. If the input system is strongly plain function passing as described in Section 10.1, then collapsing dependency pairs are dropped. If it is not strongly plain function passing, then a static approach is attempted only when the dynamic approach fails. The tool WANDA uses a simple form of argument functions immediately (if there is any constraint $f(x_1, \dots, x_n) \succeq r$ or $f(x_1, \dots, x_n) \prec_{\sim} r$ then the argument function $\pi(f(\vec{x})) = r$ is used), a proof-of-concept implementation of polynomial interpretations in the natural numbers, and our own version of a recursive path ordering (based on [22]) with argument filterings.

The current termination problem database (8.0.1) contains 156 higher-order benchmarks; out of these, 138 are local, and 133 are plain function passing. The following table sums up how the various restrictions relate.

| | local | not local | not left-linear | not fully extended |
|------------|-------|-----------|-----------------|--------------------|
| PFP | 123 | 10 | 9 | 1 |
| Strong PFP | 88 | 9 | 8 | 1 |
| Not PFP | 15 | 8 | 4 | 6 |

We executed WANDA 1.6⁴ with and without dependency pairs, and also once with the locality improvement disabled, on a dual Intel(R) Core(TM) i5-2520M CPU @ 2.50GHz computer with 4G RAM and a 60 second timeout; the results are summarised in the table below. An evaluation page with more details is available at <http://www.few.vu.nl/~femke/lmcs2012/>.

| | full WANDA | dynamic DP | static DP | rule removal | non-local |
|-----------------|------------|------------|-----------|--------------|-----------|
| YES | 122 | 119 | 114 | 75 | 100 |
| NO | 9 | 9 | 9 | 9 | 9 |
| MAYBE | 23 | 26 | 31 | 69 | 44 |
| TIMEOUT | 2 | 2 | 2 | 3 | 3 |
| average runtime | 1.33 | 1.97 | 1.40 | 0.89 | 1.82 |

In *rule removal* a reduction pair (\succeq, \succ) is used where \succ is monotonic, and for all rules either $l \succ r$ or $l \succeq r$; strictly oriented rules can be removed. The timeouts are mainly due to WANDA's strategy of trying more advanced polynomial interpretations if simple forms fail. It is noticeable that the static style is a fair bit faster than the dynamic one; this is to be expected, since it gives both fewer and easier constraints, and the subterm criterion in particular often supplies a quick solution. The results for local AFSs of Section 6 give a significant improvement over the basic dynamic dependency pair approach, but even the basic result is a real improvement over rule removal.

This table demonstrates that using a dependency pair approach, either static or dynamic, significantly improves the strength of an automatic termination tool for higher-order rewriting. Moreover, the combination of static and dynamic dependency pairs is stronger than either technique on its own; the two styles are truly incomparable.

12. CONCLUSION

We have presented a method to prove termination of AFSs using dynamic dependency pairs. A first improvement of the method is obtained, for the subclass of local AFSs, by using formative rules, which are a variation of the usable rules from the first-order case. Further improvements are obtained by the use of argument functions, which are an extension of the argument filterings from the first-order case, and the subterm criterion and usable rules for non-collapsing dependency pairs. The ordering constraints generated by the dependency pair approach can be solved not only using reduction pairs such as higher-order recursive path orderings, but also using weakly monotonic algebras.

⁴This is not the version which was used in the last termination competition; there, polynomial interpretations had not yet been implemented.

The dependency pair approach as presented here is implemented in the higher-order termination tool **WANDA** [20] by the first author. It can be used on its own, or together with a static approach in an automatic termination tool. Experiments with the termination problem database have shown that the styles are incomparable, but have similar strength.

With this work we aim to contribute to the understanding of termination of higher-order rewriting, and more in particular, to the understanding of dependency pairs, and to the developments of tools to automatically prove termination of higher-order rewriting.

For future work, it would be interesting to generalise the dependency pair *framework* [10] to the higher-order setting. This poses challenges like the question whether tags and formative rules should be flags, or can be phrased as dependency pair processors. In a different direction, we might study (and implement) the new method of formative rules in the first-order setting.

13. ACKNOWLEDGEMENTS

We gratefully acknowledge remarks and suggestions from Vincent van Oostrom and Jan Willem Klop. We thank the anonymous referees of earlier versions, and the anonymous referees of the present version for their constructive and helpful remarks.

REFERENCES

- [1] T. Aoto and Y. Yamada. Dependency pairs for simply typed term rewriting. In J. Giesl, editor, *Proceedings of RTA 2005*, volume 3467 of *LNCS*, pages 120–134, Nara, Japan, 2005. Springer.
- [2] T. Arts and J. Giesl. Termination of term rewriting using dependency pairs. *Theoretical Computer Science*, 236(1-2):133–178, 2000.
- [3] F. Blanqui. Higher-order dependency pairs. In *Proceedings of WST 2006*, Seattle, USA, 2006.
- [4] F. Blanqui, J.-P. Jouannaud, and A. Rubio. The computability path ordering: The end of a quest. In *CSL 2008*, volume 5213 of *LNCS*, pages 1–14, Bertinoro, Italy, 2008. Springer.
- [5] H.J.S. Bruggink. *Equivalence of Reductions in Higher-Order Rewriting*. PhD thesis, Utrecht University, 2008.
- [6] N. Cagman and Roger Hindley. Combinatory weak reduction in lambda-calculus. *Theoretical Computer Science*, 198:239–247, 1998.
- [7] C. Fuhs, J. Giesl, M. Parting, P. Schneider-Kamp, and S. Swiderski. Proving termination by dependency pairs and inductive theorem proving. *Journal of Automated Reasoning*, 47(2):133–160, 2011.
- [8] C. Fuhs and C. Kop. Harnessing first order termination provers using higher order dependency pairs. In C. Tinelli and V. Sofronie-Stokkermans, editors, *Frontiers of Combining Systems*, volume 6989 of *LNCS*, pages 147–162. Springer, 2011.
- [9] C. Fuhs and C. Kop. Polynomial interpretations for higher-order rewriting. In A. Tiwari, editor, *Proceedings of the 23rd International Conference on Rewriting Techniques and Applications (RTA '12)*, volume 15 of *LIPIcs*, pages 176–192. Dagstuhl, 2012.
- [10] J. Giesl, R. Thiemann, and P. Schneider-Kamp. The dependency pair framework: Combining techniques for automated termination proofs. In *Proceedings of LPAR 2004*, volume 3452 of *LNCS*, pages 301–331. Springer, 2005.
- [11] J. Giesl, R. Thiemann, P. Schneider-Kamp, and S. Falke. Mechanizing and improving dependency pairs. *Journal of Automated Reasoning*, 37(3):155–203, 2006.
- [12] B. Gramlich. Abstract relations between restricted termination and confluence properties of rewrite systems. *Fundamenta Informaticae*, 24:3–23, 1995.
- [13] N. Hirokawa and A. Middeldorp. Dependency pairs revisited. In *Proceedings of RTA 2004*, volume 3091 of *LNCS*, pages 249–268. Springer, 2004.
- [14] N. Hirokawa and A. Middeldorp. Automating the dependency pair method. *Information and Computation*, 199(1-2):172–199, 2005.
- [15] N. Hirokawa and A. Middeldorp. Tyrolean termination tool: Techniques and features. *Information and Computation*, 205(4):474–511, 2007.

- [16] N. Hirokawa, A. Middeldorp, and H. Zankl. Uncurrying for termination. In *LPAR 2008*, volume 5330 of *LNAI*, pages 667–681, Doha, 2008. Springer-Verlag.
- [17] G. Huet and D.C. Oppen. Equations and rewrite rules: a survey. In R.V. Book, editor, *Formal Language Theory: Perspectives and Open Problems*, pages 349–405. Academic Press, London, 1980.
- [18] J.-P. Jouannaud and M. Okada. A computation model for executable higher-order algebraic specification languages. In *LICS 1991*, pages 350–361, Amsterdam, The Netherlands, 1991. IEEE Computer Society Press.
- [19] J.-P. Jouannaud and A. Rubio. The higher-order recursive path ordering. In *LICS 1999*, pages 402–411, Trento, Italy, 1999.
- [20] C. Kop. Wanda. <http://www.few.vu.nl/~kop/code.html>.
- [21] C. Kop. Simplifying algebraic functional systems. In F. Winkler, editor, *Proceedings of CAI 2011*, volume 6742 of *LNCs*, pages 201–215. Springer, 2011.
- [22] C. Kop and F. van Raamsdonk. A higher-order iterative path ordering. In I. Cervesato, H. Veith, and A. Voronkov, editors, *Proceedings of LPAR 2008*, volume 5330 of *LNAI*, pages 697–711. Springer, 2008.
- [23] C. Kop and F. van Raamsdonk. Higher-order dependency pairs with argument filterings. In *Proceedings of WST 2010*, Edinburgh, UK, 2010. <http://www.few.vu.nl/~kop/wst10.pdf>.
- [24] C. Kop and F. van Raamsdonk. Higher order dependency pairs for algebraic functional systems. In M. Schmidt-Schauß, editor, *Proceedings of RTA 2011*, volume 10 of *LIPICs*, pages 203–218. Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik, 2011.
- [25] K. Kusakari. On proving termination of term rewriting systems with higher-order variables. *IPSJ Transactions on Programming*, 42(SIG 7 PRO11):35–45, 2001.
- [26] K. Kusakari, Y. Isogai, M. Sakai, and F. Blanqui. Static dependency pair method based on strong computability for higher-order rewrite systems. *IEICE Transactions on Information and Systems*, 92(10):2007–2015, 2009.
- [27] K. Kusakari and M. Sakai. Enhancing dependency pair method using strong computability in simply-typed term rewriting. *AAECC*, 18(5):407–431, 2007.
- [28] K. Kusakari and M. Sakai. Static dependency pair method for simply-typed term rewriting and related techniques. *IEICE Transactions*, 2(92-D):235–247, 2009.
- [29] P.-A. Mellès. *Description Abstraite des Systèmes de Réécriture*. PhD thesis, Université Paris VII, 1996.
- [30] T. Nipkow. Higher-order critical pairs. In *LICS 1991*, pages 342–349, Amsterdam, The Netherlands, 1991.
- [31] V. van Oostrom. Finite family developments. In H. Comon, editor, *Proceedings of RTA 1997*, volume 1232 of *LNCs*, pages 308–322, Sitges, Spain, 1997. Springer.
- [32] J.C. van de Pol. *Termination of Higher-order Rewrite Systems*. PhD thesis, University of Utrecht, 1996.
- [33] J.C. van de Pol and H. Schwichtenberg. Strict functionals for termination proofs. In Mariangiola Dezani-Ciancaglini and Gordon Plotkin, editors, *Typed Lambda Calculi and Applications*, volume 902 of *Lecture Notes in Computer Science*, pages 350–364. Springer Berlin / Heidelberg, 1995.
- [34] M. Sakai and K. Kusakari. On dependency pair method for proving termination of higher-order rewrite systems. *IEICE Transactions on Information and Systems*, E88-D(3):583–593, 2005.
- [35] M. Sakai, Y. Watanabe, and T. Sakabe. An extension of the dependency pair method for proving termination of higher-order rewrite systems. *IEICE Transactions on Information and Systems*, E84-D(8):1025–1032, 2001.
- [36] S. Suzuki, K. Kusakari, and F. Blanqui. Argument filterings and usable rules in higher-order rewrite systems. *IPSJ Transactions on Programming*, 4(2):1–12, 2011.
- [37] Terese. *Term Rewriting Systems*, volume 55 of *Cambridge Tracts in Theoretical Computer Science*. Cambridge University Press, 2003.
- [38] Wiki. Termination portal. <http://www.termination-portal.org/>.
- [39] H. Zantema. Termination of term rewriting: interpretation and type elimination. *Journal of Symbolic Computation*, 17:23–50, 1994.